

# SPARSE AND WEIGHTED ESTIMATES FOR GENERALIZED HÖRMANDER OPERATORS AND COMMUTATORS

GONZALO H. IBÁÑEZ-FIRNKORN AND ISRAEL P. RIVERA-RÍOS

**ABSTRACT.** In this paper we obtain a pointwise sparse domination for generalized Hörmander operators and also for iterated commutators with those operators. As a particular case of our result we obtain an extension of the sparse domination for commutators [24, Theorem 1.1] to iterated commutators. Relying upon that sparse domination a number of quantitative estimates such as Coifman-Fefferman estimates, strong type estimates, and endpoint estimates that improve and complete results in [30, 29, 28] are obtained. We also provide a new local decay estimate and we also extend results in [31] to kernels satisfying generalized Hörmander conditions. Among other applications, as a particular case of our result for endpoint estimates, we extend [24, Theorem 1.2] to iterated commutators.

## CONTENTS

1. Introduction and main results	2
1.1. Strong type estimates	4
1.2. Coifman-Fefferman estimates and related results	4
1.3. Endpoint estimates	5
1.4. Local exponential decay estimates	6
2. Applications	6
2.1. Weighted endpoint estimates for Coifman-Rochberg-Weiss iterated commutators	6
2.2. Some applications revisited	7
2.2.1. Homogeneous operators	7
2.2.2. Fourier Multipliers	8
3. Preliminaries	8
4. Proof of the sparse domination	13
5. Proofs of strong type estimates	18
6. Proofs of Coifman-Fefferman estimates and related results	20
7. Proofs of endpoint estimates	23
8. Proofs of exponential decay estimates	27
9. Proofs of applications	28
Appendix A. Quantitative unweighted estimates	31
Acknowledgments	32
References	32

---

2000 *Mathematics Subject Classification.* 42B20, 42B25.

*Key words and phrases.* Commutators, Generalized Hörmander conditions, Sparse operators, weighted inequalities, Calderón-Zygmund operators.

The first author is supported by CONICET and SECYT-UNC and also by the Basque Government through the BERC 2014-2017 program and by Spanish Ministry of Economy and Competitiveness MINECO: BCAM Severo Ochoa excellence accreditation SEV-2013-0323.

The second author is supported by Spanish Ministry of Economy and Competitiveness MINECO through the project MTM2014-53850-P.

## 1. INTRODUCTION AND MAIN RESULTS

During the last years a new set of techniques that allow to control operators (generally singular operators) in terms of averages over dyadic cubes has blossomed, due to fact that those kind of objects allow to simplify proofs of known results or even to obtain new results in the theory of weights. The beginning of this trend was motivated by the attempt of simplifying the original proof of the  $A_2$  Theorem [13], namely, that if  $T$  is a Calderón-Zygmund operator satisfying a Hölder-Lipschitz condition, then

$$\|Tf\|_{L^2(w)} \leq c_{n,T}[w]_{A_2} \|f\|_{L^2(w)},$$

and can be traced back to the work of A. K. Lerner [22]. In that work it is established that any standard Calderón-Zygmund operator satisfying a Hölder-Lipschitz condition can be controlled in norm by sparse operators, to be more precise, that

$$\|Tf\|_X \leq \sup_{\mathcal{S}} \|\mathcal{A}_{\mathcal{S}}f\|_X \quad (1.1)$$

where  $X$  is any Banach functions space and

$$\mathcal{A}_{\mathcal{S}}f(x) = \sum_{Q \in \mathcal{S}} \frac{1}{|Q|} \int_Q |f| \chi_Q(x)$$

where  $\mathcal{S}$  is a sparse family, namely, a family of dyadic cubes such that for each  $Q \in \mathcal{S}$  there exists a measurable  $E_Q \subseteq Q$  such that

$$|Q| \leq 2|E_Q|$$

and the  $E_Q$  are pairwise disjoint. (1.1) combined with the following estimate from [8]

$$\|\mathcal{A}_{\mathcal{S}}\|_{L^2(w) \rightarrow L^2(w)} \leq c_n[w]_{A_2}$$

yields an easy proof of the  $A_2$  Theorem. Later on it was proved independently in [6] and in [23] that

$$|Tf(x)| \leq c_n \kappa_T \sum_{j=1}^{3^n} \mathcal{A}_{\mathcal{S}_j}f(x).$$

Quite recently a fully quantitative version of this result for Calderón-Zygmund operators satisfying a Dini condition has been obtained in [19] (see [21] for a simplified proof and also [20] for the idea of the iteration technique). In that fully quantitative estimate  $\kappa_T = \|T\|_{L^2 \rightarrow L^2} + c_K + \|\omega\|_{\text{Dini}}$  where  $c_K$  denotes the size condition constant for  $T$  and  $\|\omega\|_{\text{Dini}} = \int_1^\infty \omega(t) \frac{dt}{t}$ . Such a precise control was fundamental to derive interesting results such as

$$\|T_\Omega\|_{L^2(w) \rightarrow L^2(w)} \leq c_n \|\Omega\|_{L^2(w)} [w]_{A_2}^2$$

where  $T_\Omega$  is a rough singular integral with  $\Omega \in L^\infty(\mathbb{S}^{n-1})$  (see [19]).

Sparse domination techniques have found applications among other operators such as commutators [24], rough singular integrals [5], or singular integrals satisfying an  $L^r$ -Hörmander condition [26] (see also [1]).

Let us turn our attention to that last class of operators. We say that  $T$  is an  $L^r$ -Hörmander singular operator if  $T$  is bounded on  $L^2$  and it admits the following representation

$$Tf(x) = \int_{\mathbb{R}^n} K(x, y) f(y) dy \quad (1.2)$$

provided that  $f \in \mathcal{C}_c^\infty$  and  $x \notin \text{supp } f$  where  $K : \mathbb{R}^n \times \mathbb{R}^n \setminus \{(x, x) : x \in \mathbb{R}^n\} \rightarrow \mathbb{R}$  is a locally integrable kernel satisfying the  $L^r$ -Hörmander condition, namely

$$H_{r,1} = \sup_Q \sup_{x, z \in \frac{1}{2}Q} \sum_{k=1}^{\infty} \left(2^k \cdot l(Q)\right)^n \left\| (K(x, \cdot) - K(z, \cdot)) \chi_{2^k Q \setminus 2^{k-1} Q} \right\|_{L^r, 2^k Q} < \infty.$$

$$H_{r,2} = \sup_Q \sup_{x, z \in \frac{1}{2}Q} \sum_{k=1}^{\infty} \left(2^k \cdot l(Q)\right)^n \left\| (K(\cdot, x) - K(\cdot, z)) \chi_{2^k Q \setminus 2^{k-1} Q} \right\|_{L^r, 2^k Q} < \infty.$$

As it was proved in [26],

$$|Tf(x)| \leq c_n c_T \sum_{j=1}^{3^n} \mathcal{A}_{r, \mathcal{S}_j} |f|(x)$$

where

$$\mathcal{A}_{r, \mathcal{S}} f = \sum_{Q \in \mathcal{S}} \left( \frac{1}{|Q|} \int_Q |f|^r \right)^{\frac{1}{r}} \chi_Q.$$

If we call  $\mathcal{H}_r$  the class of kernels satisfying an  $L^r$ -Hörmander condition, and  $\mathcal{H}_{\text{Dini}}$  the class of kernels satisfying a Dini condition we have that

$$\mathcal{H}_{\text{Dini}} \subset \mathcal{H}_\infty \subset \mathcal{H}_r \subset \mathcal{H}_s \subset \mathcal{H}_1 \quad 1 < s < r < \infty. \quad (1.3)$$

There's a wide range of Hörmander conditions that, somehow, lay between classes of kernels in (1.3). Those conditions based in generalizing the  $L^r$ -Hörmander condition with Young functions (cf. Subsection 3.3 for precise definitions). Let  $A$  be a Young function. We say that  $T$  is a  $A$ -Hörmander operator if  $\|T\|_{L^2 \rightarrow L^2} < \infty$ , if it satisfies a size condition and it also admits a representation as (1.2) with  $K$  belonging to the class  $\mathcal{H}_A$ , namely satisfying that  $H_A = \max \{H_{A,1}, H_{A,2}\} < \infty$  where

$$\begin{aligned} H_{A,1} &= \sup_Q \sup_{x, z \in \frac{1}{2}Q} \sum_{k=1}^{\infty} \left( 2^k \cdot l(Q) \right)^n \left\| (K(x, \cdot) - K(z, \cdot)) \chi_{2^k Q \setminus 2^{k-1} Q} \right\|_{A, 2^k Q} < \infty \\ H_{A,2} &= \sup_Q \sup_{x, z \in \frac{1}{2}Q} \sum_{k=1}^{\infty} \left( 2^k \cdot l(Q) \right)^n \left\| (K(\cdot, x) - K(\cdot, z)) \chi_{2^k Q \setminus 2^{k-1} Q} \right\|_{A, 2^k Q} < \infty. \end{aligned} \quad (1.4)$$

Operators related to that kind of conditions and commutators of  $BMO$  symbols and those operators have been thoroughly studied in several works. M. Lorente, M. S. Riveros and A. de la Torre obtained Coifman-Fefferman estimates suited for those operators [30], the same authors in a joint work with J. M. Martell established Coifman-Fefferman inequalities and also weighted endpoint estimates in the case  $w \in A_\infty$  for commutators, as well as one-sided estimates in [29]. Later on, M. Lorente, M. S. Riveros, J. M. Martell and C. Pérez proved some interesting endpoint estimates for arbitrary weights in [28]. The purpose of this work is to update and improve results in those works using sparse domination techniques.

Our first result, that will be the cornerstone for the rest of the results in this paper, is a pointwise sparse estimate for both  $A$ -Hörmander operators and commutators. We recall that given a locally integrable function  $b$  and a linear operator  $T$ , we define the commutator of  $T$  and  $b$ , by

$$[b, T]f(x) = b(x)Tf(x) - T(bf)(x).$$

We can define the iterated commutator for  $m > 1$  as

$$T_b^m f(x) = [b, T_b^{m-1}]f(x),$$

where  $T_b^1$  stands for  $[b, T]$  and making a convenient abuse of notation  $T_b^0 = T$ . Using the notation we have just introduced, we present our first result. Precise definitions of the objects and structures involved in the statement can be found in Section 3.

**Theorem 1.** *Let  $A \in \mathcal{Y}(p_0, p_1)$  a Young function with complementary function  $\bar{A}$ . Let  $T$  be an  $\bar{A}$ -Hörmander operator. Let  $m$  be a non-negative integer. For every compactly supported  $f \in C_c^\infty(\mathbb{R}^n)$  and  $b \in L_{loc}^1(\mathbb{R}^n)$ , there exist  $3^n$  dyadic lattices  $\mathcal{D}_j$  and sparse families  $\mathcal{S}_j \subseteq \mathcal{D}_j$  such that*

$$|T_b^m f(x)| \leq c_{n,m} C_T \sum_{j=1}^{3^n} \sum_{h=0}^m \binom{m}{h} \mathcal{A}_{A, \mathcal{S}_j}^{m,h}(b, f)(x)$$

where

$$\mathcal{A}_{\mathcal{S}}^{m,h}(b, f)(x) = \sum_{Q \in \mathcal{S}} |b(x) - b_Q|^{m-h} \|f|b - b_Q|^h\|_{A, Q} \chi_Q(x)$$

and  $\mathcal{A}_{A,S}^{0,0}(b, f) = \mathcal{A}_S f(x)$ .  $C_T = c_{n,p_0,p_1} \max\{c_{A,p_0}, c_{A,p_1}\} (H_{\bar{A}} + \|T\|_{L^2 \rightarrow L^2})$ .

We would like to point out that the preceding result generalizes the pointwise estimates obtained in [19, 24] since is completely new for iterated commutators and that it also provides a pointwise estimate in the case that  $T$  is a Calderón-Zygmund operator satisfying a Dini condition. Indeed, as we point out at the end of Subsection 3.3, if  $T$  is a  $\omega$ -Calderón Zygmund operator, then  $T$  is a  $L^\infty$ -Hörmander singular operator, with  $H_\infty \leq c_n(\|\omega\|_{\text{Dini}} + C_K)$  and in this case it suffices to apply our result with  $A(t) = t$  which yields the corresponding estimate with  $C_T = \|T\|_{L^2 \rightarrow L^2} + \|\omega\|_{\text{Dini}} + C_K$ .

**1.1. Strong type estimates.** Relying upon the sparse domination that we have just presented we can derive strong type quantitative estimates in terms of  $A_p - A_\infty$  constants (cf. Subsection 3.4 for precise definitions).

**Theorem 2.** *Let  $A \in \mathcal{Y}(p_0, p_1)$  be a Young function with complementary function  $\bar{A}$  and  $T$  an  $\bar{A}$ -Hörmander operator. Let  $b \in BMO$  and  $m$  be a non-negative integer. Let  $1 < p < \infty$  and  $1 < r < \infty$  and assume that  $\mathcal{K}_{r,A} = \sup_{t>1} \frac{A(t)^{\frac{1}{r}}}{t} < \infty$ . Then, for every  $w \in A_{p/r}$ ,*

$$\|T_b^m f\|_{L^p(w)} \leq c_n c_T \|b\|_{BMO}^m \mathcal{K}_{r,A} [w]_{A_{p/r}}^{\frac{1}{p}} \left( [w]_{A_\infty}^{\frac{1}{p'}} + [\sigma_{p/r}]_{A_\infty}^{\frac{1}{p}} \right) ([w]_{A_\infty} + [\sigma_{p/r}]_{A_\infty})^m \|f\|_{L^p(w)}. \quad (1.5)$$

where  $\sigma_{p/r} = w^{-\frac{1}{p'-1}}$ .

It is also possible to obtain a weighted strong type  $(p, p)$  in terms of a “bumped”  $A_p$  class.

**Theorem 3.** *Let  $B \in \mathcal{Y}(p_0, p_1)$  be a Young function with complementary function  $\bar{B}$  and let  $A, C$  be Young functions such that  $A^{-1}(t)\bar{B}^{-1}(t)C^{-1}(t) \leq ct$  for every  $t \geq t_0$  for some  $t_0$  and  $A \in B_p$ . Let  $T$  be a  $\bar{B}$ -Hörmander operator. Then if  $w \in A_p$  is a weight satisfying additionally the following condition*

$$[w]_{A_p(C)} = \sup_Q \frac{w(Q)}{|Q|} \|w^{-\frac{1}{p}}\|_{C,Q}^p < \infty$$

we have that

$$\|Tf\|_{L^p(w)} \leq c_{n,p} [w]_{A_p(C)}^{\frac{1}{p}} [w]_{A_p}^{\frac{1}{p'}} \|f\|_{L^p(w)}. \quad (1.6)$$

Even though Theorems 2 and 3 provide interesting quantitative weighted estimates, it would be desirable, if it is possible, to obtain some result in terms of some bump condition suited for each class of kernels  $\mathcal{H}_{\bar{A}}$  that reduces to the standard  $A_{p/r}$  condition for the class  $\mathcal{H}_{r'}$ .

**1.2. Coifman-Fefferman estimates and related results.** Now we turn our attention to Coifman-Fefferman type estimates. We obtain the following result.

**Theorem 4.** *Let  $B$  be a Young function such that  $B \in \mathcal{Y}(p_0, p_1)$ . If  $T$  is a  $\bar{B}$ -Hörmander operator, then for any  $1 \leq p < \infty$  and any weight  $w \in A_\infty$ ,*

$$\|Tf\|_{L^p(w)} \leq c_n [w]_{A_\infty} \|M_B f\|_{L^p(w)} \quad (1.7)$$

*If additionally  $b \in BMO$ ,  $m$  is a non-negative integer and  $A$  is a Young function, such that  $A^{-1}(t)\bar{B}^{-1}(t)\bar{C}^{-1}(t) \leq t$  with  $\bar{C}(t) = e^{t^{1/m}}$  for  $t \geq 1$ , then for any  $1 \leq p < \infty$  and any weight  $w \in A_\infty$*

$$\|T_b^m f\|_{L^p(w)} \leq c_{n,m} \|b\|_{BMO}^m [w]_{A_\infty}^{m+1} \|M_A f\|_{L^p(w)} \quad (1.8)$$

We would like to point out that Theorem 4 was proved in [30] for operators satisfying an  $A$ -Hörmander condition. Later on in [29, Theorem 3.3] a suitable version of this estimate for commutators was also obtained. Theorem 4 improves the results in [30, 29] in two directions. It provides quantitative estimates for the range  $1 \leq p < \infty$  and in the case  $m > 0$  the class of operators considered is also wider. This estimate can be extended to the full range  $0 < p < \infty$  using Rubio de Francia extrapolation arguments in [7, 9] but without a precise control of the dependence on the

$A_\infty$  constant. We encourage the reader to consult them to gain a profound insight into Rubio de Francia extrapolation techniques and the results that can be obtained from them.

Related to the sharpness of the preceding result, in [31] it was established that  $L^r$ -Hörmander condition is not enough for a convolution type operator to have a full weight theory. In the following Theorem we extend that result to a certain family of  $A$ -Hörmander operators.

**Theorem 5.** *Let  $1 \leq r < \infty$ ,  $1 \leq p < r'$  and  $\frac{p}{r'} < \gamma < 1$ . Let  $A$  a Young function such that*

$$A^{-1}(t) \simeq \frac{t^{\frac{1}{r}}}{\varphi(t)} \quad \text{for } t > c_A > 0$$

*where  $0 < \varphi(t) < \kappa_s t^s$  for every  $t > c_s > 0$  and  $1 \leq s < \infty$ . Then there exists an operator  $T$  satisfying an  $A$ -Hörmander condition such that*

$$\|T\|_{L^p(w) \rightarrow L^{p,\infty}(w)} = \infty$$

*where  $w(x) = |x|^{-\gamma n}$ .*

From this result, via extrapolation techniques, it also follows, using ideas in [31] that the Coifman-Fefferman estimate 1.7, does not hold for maximal operators that are not big enough.

**Theorem 6.** *Let  $1 \leq r < \infty$ . Let  $A$  a Young function such that*

$$A^{-1}(t) \simeq \frac{t^{\frac{1}{r}}}{\varphi(t)} \quad \text{for } t > c_A > 0$$

*where  $0 < \varphi(t) < \kappa_s t^s$  for every  $t > c_s > 0$  and  $1 \leq s < \infty$ . There exists an operator  $T$  satisfying an  $A$ -Hörmander condition such that for each  $1 < q < r'$  and  $B(t) \leq ct^q$ , the following estimate*

$$\|Tf\|_{L^p(w)} \leq c \|M_B f\|_{L^p(w)} \quad (1.9)$$

*where  $w \in A_\infty$  does not hold for any  $0 < p < \infty$  and any constant  $c$  depending on  $w$ .*

**1.3. Endpoint estimates.** In this subsection we present some quantitative endpoint estimates that can be obtained following ideas in [10, 24]. For the sake of clarity in this case we will present different statements for  $T$  and  $T_b^m$  with  $m$  a positive integer.

**Theorem 7.** *Let  $A \in \mathcal{Y}(p_0, p_1)$  a Young function and  $T$  an  $\bar{A}$ -Hörmander operator. Assume that  $A$  is submultiplicative, namely, that  $A(xy) \leq A(x)A(y)$ . Then we have that for every weight  $w$ , and every Young function  $\varphi$ ,*

$$w(\{x \in \mathbb{R}^n : Tf(x) > \lambda\}) \leq c_n C_T \kappa_\varphi \int_{\mathbb{R}^n} A\left(\frac{|f(x)|}{\lambda}\right) M_\varphi w(x) dx, \quad (1.10)$$

*where*

$$\kappa_\varphi = \int_1^\infty \frac{\varphi^{-1}(t) A(\log(e+t)^2)}{t^2 \log(e+t)^3} dt.$$

For commutators we have the following result.

**Theorem 8.** *Let  $b \in BMO$  and  $m$  be a positive integer. Let  $A_0, \dots, A_m$  be Young functions, such that  $A_0 \in \mathcal{Y}(p_0, p_1)$  and  $A_j^{-1}(t) \bar{A}_0^{-1}(t) \bar{C}_j^{-1}(t) \leq t$  with  $\bar{C}_j(t) = e^{t^{\frac{1}{j}}}$  for  $t \geq 1$ . Let  $T$  be a  $\bar{A}_0$ -Hörmander operator. Assume that each  $A_j$  is submultiplicative, namely, that  $A_j(xy) \leq A_j(x)A_j(y)$ . Then we have that for every weight  $w$ , and every family of Young functions  $\varphi_0, \dots, \varphi_m$*

$$w(\{x \in \mathbb{R}^n : T_b^m f(x) > \lambda\}) \leq c_n C_T \sum_{h=0}^m \left( \kappa_{\varphi_h} \int_{\mathbb{R}^n} A_h\left(\frac{|f(x)|}{\lambda}\right) M_{\Phi_{m-h} \circ \varphi_h} w(x) dx \right), \quad (1.11)$$

*where  $\Phi_j(t) = t \log(e+t)^j$ ,  $0 \leq j \leq m$ ,*

$$\kappa_{\varphi_h} = \begin{cases} \alpha_{n,m,h} + c_n \int_1^\infty \frac{\varphi_h^{-1} \circ \Phi_{m-h}^{-1}(t) A_h(\log(e+t)^{4(m-h)})}{t^2 \log(e+t)^{3(m-h)+1}} dt & 0 \leq h < m, \\ \int_1^\infty \frac{\varphi_h^{-1}(t) A_h(\log(e+t)^2)}{t^2 \log(e+t)^3} dt & h = m. \end{cases}$$

At this point we would like to make some remarks about Theorems 7 and 8. These results provide quantitative versions of [29, Theorem 3.3] for arbitrary weights instead of considering just  $A_\infty$  weights. We also recall that in the case of  $T$  satisfying an  $A$ -Hörmander condition, it is proved in [28, Theorem 3.1] that  $T$  satisfies a weak-type  $(1, 1)$  inequality for a pair of weights  $(u, Su)$  where  $S$  is a suitable maximal operator. We observe that it is not possible to recover  $A_1$  estimates from those results, since otherwise that would lead to a contradiction with [31, Theorem 3.2] or with Theorem 5. Hence Theorem 7 and [28, Theorem 3.1] are complementary results. Theorems 8 and [28, Theorem 3.8] could be compared in an analogous way.

In Subsection 2.1 we will present an application of Theorem 8 to the case in which  $T$  an  $\omega$ -Calderón-Zygmund operator that provides a new weighted endpoint for iterated commutators that extends naturally [24, Theorem 1.2].

**1.4. Local exponential decay estimates.** Also as a consequence of the sparse domination result we can derive the following local estimates, in the spirit of [34].

**Theorem 9.** *Let  $B$  a Young function such that  $B \in \mathcal{Y}(p_0, p_1)$  and  $T$  a  $\bar{B}$ -Hörmander operator. Let  $f$  be a function such that  $\text{supp } f \subseteq Q$ . Then there exist constants  $c_n$  and  $\alpha_n$  such that*

$$\left| \left\{ x \in Q : \frac{|Tf(x)|}{M_B f(x)} > \lambda \right\} \right| \leq c_n e^{-\alpha_n \frac{\lambda}{c_T}} |Q|. \quad (1.12)$$

*If additionally  $m$  is a positive integer,  $b \in BMO$  and  $A$  is a Young function that satisfies the following inequality  $A^{-1}(t)\bar{B}^{-1}(t)\bar{C}^{-1}(t) \leq t$  with  $\bar{C}(t) = e^{t^{1/m}}$  for  $t \geq 1$ , then there exist constants  $c_{n,m}$  and  $\alpha_{n,m}$  such that*

$$\left| \left\{ x \in Q : \frac{|T_b^m f(x)|}{M_A f(x)} > \lambda \right\} \right| \leq c_{n,m} e^{-\alpha_{n,m} \left( \frac{\lambda}{c_T \|b\|_{BMO}^m} \right)^{\frac{1}{m+1}}} |Q| \quad \lambda > 0. \quad (1.13)$$

## 2. APPLICATIONS

In this section we gather some applications of the main theorems. We present an extension of [24, Theorem 1.2] to iterated commutators, which is completely new. We also revisit some applications that appeared in [29].

### 2.1. Weighted endpoint estimates for Coifman-Rochberg-Weiss iterated commutators.

R. Coifman, R. Rochberg and G. Weiss introduced the commutator of a Calderón-Zygmund operator with a  $BMO$  symbol in [4] to study the factorization of  $n$ -dimensional Hardy spaces. Those commutators were proved to not to be of weak type  $(1, 1)$  in [36] where a suitable endpoint replacement for them and for iterated commutators as well, namely a distributional estimate, was also provided for Lebesgue measure and  $A_1$  weights.

In [37] C. Pérez and G. Pradolini obtained an endpoint estimate for commutators with arbitrary weights, and later on, C. Pérez and the second author [38] obtained a quantitative version of that result that reads as follows

$$w(\{x \in \mathbb{R}^n : |T_b^m f(x)| > \lambda\}) \leq c \frac{1}{\varepsilon^{m+1}} \int_{\mathbb{R}^n} \Phi_m \left( \frac{|f|}{\lambda} \right) M_{L(\log L)^{m+\varepsilon}} w \quad \varepsilon > 0,$$

where  $\Phi_m(t) = t \log(e + t)^m$ . From that estimate is possible to recover the following estimates that are essentially contained in [33]

$$\begin{aligned} w(\{x \in \mathbb{R}^n : |T_b^m f(x)| > \lambda\}) &\leq c[w]_{A_\infty}^m \log(e + [w]_{A_\infty})^{m+1} \int_{\mathbb{R}^n} \Phi_m \left( \frac{|f|}{\lambda} \right) Mw \quad w \in A_\infty \\ &\leq c[w]_{A_1} [w]_{A_\infty}^m \log(e + [w]_{A_\infty})^{m+1} \int_{\mathbb{R}^n} \Phi_m \left( \frac{|f|}{\lambda} \right) w \quad w \in A_1 \end{aligned}$$

In the case  $m = 1$  it was established in [24] that the blow up in  $\frac{1}{\varepsilon}$  is linear instead of quadratic. That improvement on the blow up led to a logarithmic improvement on the dependence on the  $A_\infty$  constant, namely,

$$\begin{aligned} w(\{x \in \mathbb{R}^n : |[b, T]f(x)| > \lambda\}) &\leq c[w]_{A_\infty} \log(e + [w]_{A_\infty}) \int_{\mathbb{R}^n} \Phi\left(\frac{|f|}{\lambda}\right) Mw \quad w \in A_\infty \\ &\leq c[w]_{A_1} [w]_{A_\infty} \log(e + [w]_{A_\infty}) \int_{\mathbb{R}^n} \Phi\left(\frac{|f|}{\lambda}\right) w \quad w \in A_1. \end{aligned}$$

In the following result we show that the same linear blow up is satisfied in the case of the iterated commutator.

**Theorem 10.** *Let  $T$  be a  $\omega$ -Calderón-Zygmund operator with  $\omega$  satisfying a Dini condition. Let  $m$  be a non-negative integer and  $b \in BMO$ . Then we have that for every weight  $w$  and every  $\varepsilon > 0$ ,*

$$\begin{aligned} w(\{x \in \mathbb{R}^n : |T_b^m f(x)| > \lambda\}) &\leq c_{n,m} C_T \frac{1}{\varepsilon} \int_{\mathbb{R}^n} \Phi_m\left(\frac{|f(x)|}{\lambda}\right) M_{L(\log L)^m (\log \log L)^{1+\varepsilon}} w(x) dx \\ &\leq c_{n,m} C_T \frac{1}{\varepsilon} \int_{\mathbb{R}^n} \Phi_m\left(\frac{|f(x)|}{\lambda}\right) M_{L(\log L)^{m+\varepsilon}} w(x) dx \end{aligned} \quad (2.1)$$

where  $\Phi_m(t) = t \log(e + t)^m$  and  $C_T = C_K + \|T\|_{L^2 \rightarrow L^2} + \|\omega\|_{Dini}$ . If additionally  $w \in A_\infty$  then

$$w(\{x \in \mathbb{R}^n : T_b^m f(x) > \lambda\}) \leq c_{n,m} C_T [w]_{A_\infty}^m \log(e + [w]_{A_\infty}) \int_{\mathbb{R}^n} \Phi_m\left(\frac{|f(x)|}{\lambda}\right) Mw(x) dx. \quad (2.2)$$

Furthermore if  $w \in A_1$

$$w(\{x \in \mathbb{R}^n : T_b^m f(x) > \lambda\}) \leq c_{n,m} C_T [w]_{A_1} [w]_{A_\infty}^m \log(e + [w]_{A_\infty}) \int_{\mathbb{R}^n} \Phi_m\left(\frac{|f(x)|}{\lambda}\right) w(x) dx. \quad (2.3)$$

We observe that Theorem 10 improves known estimates in two directions. We improve the maximal operator that we need in the right hand side of the estimate for it to hold, and the blow up in  $\frac{1}{\varepsilon}$ , which leads to a logarithmic improvement of the dependence on the  $A_\infty$  constant.

**2.2. Some applications revisited.** In this section we revisit some applications that were covered in [29], namely Homogeneous operators and Fourier Multipliers. We improve known results for those operators and we also provide some new estimates.

**2.2.1. Homogeneous operators.** Let  $\Omega \in L^1(\mathbb{S}^{n-1})$  such that  $\int_{\mathbb{S}^{n-1}} \Omega = 0$ . Setting  $K(x) = \frac{\Omega(x)}{|x|^n}$  we consider the following convolution type operator

$$T_\Omega f(x) = \text{p.v.} \int_{\mathbb{R}^n} K(x-y) f(y) dy.$$

Our result is the following.

**Theorem 11.** *Let  $T_\Omega$  be as above. Let  $B$  a Young function such that  $B \in \mathcal{Y}(p_0, p_1)$  and*

$$\int_0^1 \omega_{\overline{B}}(t) \frac{dt}{t} < \infty \quad (2.4)$$

where

$$\omega_{\overline{B}}(t) = \sup_{|y| \leq t} \|\Omega(\cdot + y) - \Omega(y)\|_{\overline{B}, \mathbb{S}^{n-1}}$$

Then  $K \in \mathcal{H}_{\overline{B}}$ . Assume that  $B \in \mathcal{Y}(p_0, p_1)$ . Then we have that

- (1) (1.7), (1.12), (1.6) and (1.10) hold for  $T_\Omega$ .
- (2) If  $m$  is a non-negative integer and  $b \in BMO$ , (1.5) holds for every  $p > r$  such that  $\mathcal{K}_{r,B} < \infty$ .
- (3) If there exists a Young function  $A$  such that  $A^{-1}(t) \overline{B}^{-1}(t) \overline{C}_m^{-1}(t) \leq t$  for every  $t \geq 1$  where  $\overline{C}_m(t) = e^{t^{1/m}}$  with  $m$  a positive integer, and  $b \in BMO$ , then we have that 1.8, (1.11) and (1.13) hold for  $(T_\Omega)_b^m$ .

This result improves and extends [29, Theorem 4.1] since we impose a weaker condition on  $\overline{B}$  and we obtain quantitative estimates and a local exponential decay estimate that are new for this operator.

**2.2.2. Fourier Multipliers.** Let  $h \in L^\infty$  we can consider a multiplier operator  $T$  defined for  $f \in \mathcal{S}$ , the Schwartz space, by

$$\widehat{Tf}(\xi) = h(\xi)\hat{f}(\xi).$$

Given  $1 < s \leq 2$  and  $l$  a non-negative integer, we say that  $h \in M(s, l)$  if

$$\sup_{R>0} R^{|\alpha|} \|D^\alpha h\|_{L^s, Q(0, 2R) \setminus Q(0, R)} < \infty$$

for all  $|\alpha| \leq l$ .

The following Coifman-Fefferman estimate was obtained in [29, Theorem 4.5].

**Theorem 12.** *Let  $h \in M(s, l)$  with  $1 < s \leq 2$ ,  $0 \leq l \leq n$  and  $l > \frac{n}{s}$ . Then for all non-negative integer  $m$  and any  $\varepsilon > 0$  we have that for all  $0 < p < \infty$  and  $w \in A_\infty$*

$$\int_{\mathbb{R}^n} |T_b^m f(x)|^p w(x) dx \leq c_{n,p,A_\infty} \int_{\mathbb{R}^n} M_{n/l+\varepsilon} f(x)^p w(x) dx.$$

The proof of that result relies upon the fact that certain truncations  $K^N$  of the kernel belong to the class  $\mathcal{H}_A$  [29, Proposition 6.2]. Here we state a slightly weaker version of their result that is enough for our purposes.

**Lemma 1.** *Let  $h \in M(s, l)$  with  $1 < s \leq 2$ ,  $1 \leq l \leq n$  and with  $l > \frac{n}{s}$ , then for every non-negative integer  $m$  and all  $1 < r < \left(\frac{n}{l}\right)'$  we have that  $K^N \in \mathcal{H}_{L^r(\log L)^{mr}}$  uniformly in  $N$ .*

Relying upon this Lemma we can obtain the following result which turns out to be new.

**Theorem 13.** *Let  $h \in M(s, l)$  with  $1 < s \leq 2$ ,  $1 \leq l \leq n$  and with  $l > \frac{n}{s}$ . Let  $m$  a non-negative integer and  $b \in BMO$ . Then,*

- (1) (1.7) and (1.8) hold with  $A(t) = t^{\frac{n}{l}+\varepsilon}$ .
- (2) If  $p > \frac{n}{l} + \varepsilon$  we have that

$$\|T_b^m f\|_{L^p(w)} \leq c_n \|b\|_{BMO}^m [w]_{A_\infty}^{\frac{1}{p}} \left( [w]_{A_\infty}^{\frac{1}{p'}} + [\sigma]_{A_\infty}^{\frac{1}{p}} \right) ([w]_{A_\infty} + [\sigma]_{A_\infty})^m \|f\|_{L^p(w)}$$

for every  $w \in A_{\frac{p}{\frac{n}{l}+\varepsilon}}$ .

### 3. PRELIMINARIES

**3.1. Dyadic structures.** In this subsection we are going to make a brief overview of the ideas of dyadic lattice, sparse family and some related results that we borrow from [23]. We encourage the reader to consult that reference for a thoroughly detailed study of that kind of structures.

We call  $\mathcal{D}(Q)$  the dyadic grid obtained repeatedly subdividing  $Q$  and its descendants in  $2^n$  cubes.

**Definition 1.** A dyadic lattice  $\mathcal{D}$  in  $\mathbb{R}^n$  is a family of cubes that satisfies the following properties

- (1) If  $Q \in \mathcal{D}$  then each descendant of  $Q$  is in  $\mathcal{D}$  as well.
- (2) For every pair of cubes  $Q_1, Q_2$  we can find a common ancestor, that is, a cube  $Q \in \mathcal{D}$  such that  $Q_1, Q_2 \in \mathcal{D}(Q)$ .
- (3) For every compact set  $K$  there exists a cube  $Q \in \mathcal{D}$  such that  $K \subseteq Q$ .

Now that we have the definition of dyadic lattice we can recall the definition of sparse family.

**Definition 2.** Let  $\mathcal{D}$  a dyadic lattice.  $\mathcal{S} \subseteq \mathcal{D}$  is a  $\eta$ -sparse family with  $\eta \in (0, 1)$  if for each  $Q \in \mathcal{S}$  we can find a measurable subset  $E_Q \subseteq Q$  such that

$$\eta|Q| \leq |E_Q|$$

and all the  $E_Q$  are pairwise disjoint.



We end this section recalling the so called  $3^n$ -dyadic lattices trick.

**Lemma 2.** *Given a dyadic lattice  $\mathcal{D}$  there exist  $3^n$  dyadic lattices  $\mathcal{D}_j$  such that*

$$\{3Q : Q \in \mathcal{D}\} = \bigcup_{j=1}^{3^n} \mathcal{D}_j$$

and for every cube  $Q \in \mathcal{D}$  we can find a cube  $R_Q$  in each  $\mathcal{D}_j$  such that  $Q \subseteq R_Q$  and  $3l_Q = l_{R_Q}$

*Remark 1.* Let us fix a dyadic lattice  $\mathcal{D}$ . For an arbitrary cube  $Q \subseteq \mathbb{R}^n$  we can find a cube  $Q' \in \mathcal{D}$  such that  $\frac{l_Q}{2} < l_{Q'} \leq l_Q$  and  $Q \subseteq 3Q'$ . It suffices to take the cube  $Q'$  that contains the center of  $Q$ . From the preceding lemma it follows that  $3Q' = P \in \mathcal{D}_j$  for some  $j \in \{1, \dots, 3^n\}$ . Therefore, for every cube  $Q \subseteq \mathbb{R}^n$  there exists  $P \in \mathcal{D}_j$  such that  $Q \subseteq P$  and  $l_P \leq 3l_Q$ . From this follows that  $|Q| \leq |P| \leq 3^n|Q|$

**3.2. Young functions and Orlicz spaces.** In this Subsection we present some notions about Young functions and Orlicz local averages that will be fundamental throughout all this work. We will not go into details for any of the results and definitions we review here. The interested reader can get profound insight into this topic in classical references such as [32], [40].

A function  $A : [0, \infty) \rightarrow [0, \infty)$  is said to be a Young function if  $A$  is continuous, convex, and satisfies that  $A(0) = 0$ . Since  $A$  is convex, we have also that  $\frac{A(t)}{t}$  is not decreasing.

We define the class of functions  $\mathcal{Y}(p_0, p_1)$  with  $1 \leq p_0 \leq p_1 < \infty$  as the class of functions  $A$  for which there exist constants  $c_{A,p_0}, c_{A,p_1}, t_A \geq 1$  such that  $t^{p_0} \leq c_{A,p_0}A(t)$  for every  $t > t_A$  and  $t^{p_1} \leq c_{A,p_1}A(t)$  for every  $t \leq t_A$ .

The average of the Luxemburg norm of a function  $f$  induced for a Young function  $A$  in the cube  $Q$  is defined by

$$\|f\|_{A(\mu),Q} := \inf \left\{ \lambda > 0 : \frac{1}{\mu(Q)} \int_Q A \left( \frac{|f|}{\lambda} \right) d\mu \leq 1 \right\} \quad (3.1)$$

If we consider  $\mu$  to be the Lebesgue measure we will write just  $\|f\|_{A,Q}$  and if  $\mu = wdx$  is an absolutely continuous measure with respect to the Lebesgue measure we will write  $\|f\|_{A(w),Q}$ .

There are several interesting facts that we review now. First we would like to note that if  $A(t) = t^r$ ,  $r \geq 1$ , then  $\|f\|_{A,Q} = \left( \frac{1}{|Q|} \int_Q |f|^r \right)^{1/r}$ , that is, we recover the standard  $L^r \left( Q, \frac{dx}{|Q|} \right)$  norm.

Another interesting fact is the following.

**Proposition 1.** *If  $A, B$  are Young function such that  $A(t) \leq \kappa B(t)$  for all  $t \geq c$ , then*

$$\|f\|_{A(\mu),Q} \leq (A(c) + \kappa) \|f\|_{B(\mu),Q}$$

for every cube  $Q$ .

We observe that the convexity of  $A$  implies that  $A(t)/t$  is not decreasing and so  $t \leq cA(t)$  for all  $t \geq 1$ , then  $\|f\|_{L^1,Q} \leq (A(1) + c) \|f\|_{A,Q}$ .

For each Young function  $A$  it is possible to define a function  $\bar{A}$ , called the complementary of  $A$ , as follows

$$\bar{A}(t) = \sup_{s>0} \{st - A(s)\}.$$

$\bar{A}$  is also a Young function and it satisfies some interesting properties such as a generalized Hölder inequality

$$\frac{1}{\mu(Q)} \int_Q |fg| d\mu \leq 2 \|f\|_{A(\mu),Q} \|g\|_{\bar{A}(\mu),Q}. \quad (3.2)$$

It also satisfies the following estimate that will be useful for us

$$t \leq A^{-1}(t) \bar{A}^{-1}(t) \leq 2t \quad (3.3)$$

and it can be proved that  $\bar{\bar{A}} \simeq A$ .

It is possible to obtain more general versions of Hölder inequality. If  $A$  and  $B$  are strictly increasing functions and  $C$  is Young such that  $A^{-1}(t)B^{-1}(t)C^{-1}(t) \leq t$ , for all  $t \geq 1$ , then

$$\|fg\|_{\bar{C}(\mu),Q} \leq c\|f\|_{A(\mu),Q}\|g\|_{B(\mu),Q}. \quad (3.4)$$

Now we turn our attention to a particular case that will be useful for us. If  $B$  is a Young function and  $A$  is a strictly increasing function such that  $A^{-1}(t)\bar{B}^{-1}(t)C^{-1}(t) \leq t$  with  $C^{-1}(t) = e^{t^{1/m}}$  for  $t \geq 1$ , then,

$$\|fg\|_{B(\mu),Q} \leq c\|f\|_{\exp L^{1/m}(\mu),Q}\|g\|_{A(\mu),Q} \leq c\|f\|_{\exp L^{1/h}(\mu),Q}\|g\|_{A(\mu),Q} \quad (3.5)$$

for all  $1 \leq h \leq m$ .

The averages that we have presented in (3.1) lead to define new maximal operators in a very natural way. Given  $f \in L^1_{\text{loc}}(\mathbb{R}^n)$ , the maximal operator associated to the Young function  $A$  is defined as

$$M_A f(x) := \sup_{Q \ni x} \|f\|_{A,Q}.$$

This kind of maximal operator was thoroughly studied in [35]. There it was established that if  $A$  is doubling and  $A \in B_p$ , namely if

$$\int_1^\infty \frac{A(t)}{t^p} \frac{dt}{t} < \infty,$$

then  $\|M_A\|_{L^p} < \infty$ . Later on T. Luque and L. Liu [27], proved that imposing the doubling condition on  $A$  is superfluous.

Now we compile some examples of maximal operators related to certain Young functions.

- $A(t) = t^r$  with  $1 < r < \infty$ . In that case  $\bar{A}(t) \simeq t^{r'}$  with  $\frac{1}{r} + \frac{1}{r'} = 1$  and then  $M_A = M_r$ .
- $A(t) = t \log(e+t)^\alpha$  with  $\alpha > 0$ . Then  $\bar{A}(t) \simeq e^{t^{1/\alpha}} - 1$ ,  $M_A = M_{L \log L^\alpha}$ . We observe that  $0 < \alpha$ , then  $M \lesssim M_A \lesssim M_r$  for all  $1 < r < \infty$ , and if  $\alpha = l \in \mathbb{N}$  it can be proved that  $M_A \approx M^{l+1}$ , where  $M^{l+1}$  is  $M$  iterated  $l+1$  times.
- If we consider  $A(t) = t \log(e+t)^l \log(e+\log(e+t))^\alpha$  with  $l, \alpha > 0$ , then we will denote  $M_A = M_{L(\log L)^l(\log \log L)^\alpha}$ . We observe that

$$M_{L(\log L)^m(\log \log L)^{1+\varepsilon}} w \leq c_\varepsilon M_{L(\log L)^{m+\varepsilon}} w \quad \varepsilon > 0.$$

We end this subsection recalling a Fefferman-Stein estimate suited for  $M_A$  that we borrow from [24, Lemma 2.6].

**Lemma 3.** *Let  $A$  be a Young function. For any arbitrary weight  $w$  we have that*

$$w(\{x \in \mathbb{R}^n : M_A f(x) > \lambda\}) \leq 3^n \int_{\mathbb{R}^n} A\left(\frac{9^n |f(x)|}{\lambda}\right) M w(x) dx.$$

*If additionally  $A$  is submultiplicative, namely  $A(xy) \leq A(x)A(y)$  then*

$$w(\{x \in \mathbb{R}^n : M_A f(x) > \lambda\}) \leq c_n \int_{\mathbb{R}^n} A\left(\frac{|f(x)|}{\lambda}\right) M w(x) dx.$$

**3.3. Singular operators.** We say that  $T$  is a singular integral operator if  $T$  is bounded on  $L^2$  and it admits the following representation

$$Tf(x) = \int_{\mathbb{R}^n} K(x,y)f(y)dy \quad \text{for all } x \notin \text{supp } f$$

where  $f \in L^1_{\text{loc}}(\mathbb{R}^n)$  and  $K : \mathbb{R}^n \times \mathbb{R}^n \setminus \{(x,x) : x \in \mathbb{R}^n\} \rightarrow \mathbb{R}$  is a locally integrable kernel away of the diagonal such that  $K \in \mathcal{H}$  for some class  $\mathcal{H}$ . Among the classes we consider in this work we recall that  $K \in \mathcal{H}_{\text{Dini}}$  if besides satisfying all the properties above,  $K$  also satisfies the a size condition

$$|K(x,y)| \leq \frac{c_K}{|x-y|^n}$$

and a smoothness condition

$$|K(x, y) - K(x', y)| + |K(y, x) - K(y, x')| \leq \omega \left( \frac{|x - x'|}{|x - y|} \right) \frac{1}{|x - y|^n},$$

for  $|x - y| > 2|x - x'|$ , where  $\omega : [0, 1] \rightarrow [0, \infty)$  is a modulus of continuity, that is a continuous, increasing, submultiplicative function with  $\omega(0) = 0$  and such that it satisfies the Dini condition, namely

$$\|\omega\|_{\text{Dini}} = \int_0^1 \omega(t) \frac{dt}{t} < \infty.$$

In this case, following the standard terminology, we say that  $T$  is a  $\omega$ -Calderón-Zygmund operator. We note that if we choose  $\omega(t) = ct^\delta$  for any  $\delta > 0$  we recover the standard Hölder-Lipschitz condition.

Given a Young function  $A$ , we say that  $K \in \mathcal{H}_A$  if  $K$  is an  $A$ -Hörmander kernel, namely if as we established in (1.4)  $H_A = \max\{H_{A,1}, H_{A,2}\} < \infty$  with

$$H_{A,1} = \sup_Q \sup_{x, z \in \frac{1}{2}Q} \sum_{k=1}^{\infty} \left( 2^k \cdot l(Q) \right)^n \left\| (K(x, \cdot) - K(z, \cdot)) \chi_{2^k Q \setminus 2^{k-1} Q} \right\|_{A, 2^k Q} < \infty,$$

$$H_{A,2} = \sup_Q \sup_{x, z \in \frac{1}{2}Q} \sum_{k=1}^{\infty} \left( 2^k \cdot l(Q) \right)^n \left\| (K(\cdot, x) - K(\cdot, z)) \chi_{2^k Q \setminus 2^{k-1} Q} \right\|_{A, 2^k Q} < \infty,$$

where the supremums are taken over all cubes  $Q$  in  $\mathbb{R}^n$ . In this case we say that a singular operator  $T$  satisfying this condition is an  $A$ -Hörmander operator. As we did in the introduction, we denote by  $\mathcal{H}_r$  the class corresponding to  $A(t) = ct^r$  for  $1 \leq r < \infty$ . Analogously we say that  $K \in \mathcal{H}_\infty$  if  $K$  satisfies the previous conditions with  $\|\cdot\|_{L^\infty, 2^k Q}$  in place of  $\|\cdot\|_{A, 2^k Q}$ . Abusing of notation, we would like to point out that if we consider  $A(t) = t$ , then

$$\overline{A}(t) = \sup_{s>0} \{st - A(s)\} = \sup_{s>0} \{(t-1)s\} = \begin{cases} 0 & t \leq 1 \\ \infty & t > 1 \end{cases}$$

so we may assume in that case that  $\overline{A}(t) = \infty$ .

It's straightforward that equivalent conditions can be stated in terms of balls instead of cubes.

Now we observe that taking into account Proposition 1, if  $A$  and  $B$  are Young functions such that there exists some  $t_0$  such that  $A(t) \leq \kappa B(t)$  every for every  $t > t_0$ , then

$$\mathcal{H}_B \subset \mathcal{H}_A.$$

Taking that property into account it is clear that the relations between the different classes of kernels presented in (1.3) hold. In particular we would like to stress the fact that if  $K \in \mathcal{H}_{\text{Dini}}$  then  $K \in \mathcal{H}_\infty$  with  $H_\infty \leq c_n(\|\omega\|_{\text{Dini}} + c_K)$ .

**3.4.  $A_p$  weights and BMO.** A function  $w$  is a weight if  $w \geq 0$  and  $w$  is locally integrable in  $\mathbb{R}^n$ . We recall that the  $A_p$  class  $1 < p < \infty$  is the class of weights  $w$  such that

$$[w]_{A_p} := \sup_Q \left( \frac{1}{|Q|} \int_Q w \right) \left( \frac{1}{|Q|} \int_Q w^{-\frac{1}{p-1}} \right)^{p-1} < \infty.$$

where the supremum is taken over all cubes  $Q$  in  $\mathbb{R}^n$ . For  $p = 1$ ,  $w \in A_1$  if and only if

$$[w]_{A_1} = \text{ess sup}_{x \in \mathbb{R}^n} \frac{Mw(x)}{w(x)} < \infty.$$

The importance of those classes of weights stems from the fact that they characterize the weighted strong-type  $(p, p)$  estimate of the Hardy-Littlewood maximal operator for  $p > 1$  and the weighted weak-type  $(1, 1)$  in the case  $p = 1$ . We observe that among other properties those classes are increasing, namely, so it is natural to define an  $A_\infty$  class as follows

$$A_\infty = \bigcup_{p \geq 1} A_p.$$

It is possible to characterize the  $A_\infty$  class in terms of a constant. In particular, it was essentially proved by Fujii [12] and later on rediscovered by Wilson [41] that

$$w \in A_\infty \iff [w]_{A_\infty} = \sup_Q \frac{1}{w(Q)} \int_Q M(w\chi_Q) < \infty.$$

In [16] this  $A_\infty$  constant was proved to be the most suitable one and the following Reverse Hölder inequality was also obtained (see [18] for another proof).

**Lemma 4.** *Let  $w \in A_\infty$ . Then for every cube  $Q$ ,*

$$\left( \frac{1}{|Q|} \int_Q w^r \right)^{\frac{1}{r}} \leq \frac{2}{|Q|} \int_Q w$$

where  $1 \leq r \leq 1 + \frac{1}{\tau_n[w]_{A_\infty}}$  with  $\tau_n$  a dimensional constant independent  $w$  and  $Q$ .

Reverse Hölder inequality allows us to give a quantitative version of one of the classical characterizations of  $A_\infty$  weights [25].

**Lemma 5.** *There exists  $c_n > 0$  such that for every  $w \in A_\infty$ , every cube  $Q$  and every measurable subset  $E \subset Q$  we have that*

$$\frac{w(E)}{w(Q)} \leq 2 \left( \frac{|E|}{|Q|} \right)^{\frac{1}{c_n[w]_{A_\infty}}}$$

*Proof.* Let us call  $r_w = 1 + \frac{1}{\tau_n[w]_{A_\infty}}$  where  $\tau_n$  is the same as in Lemma (4). We observe that using Reverse Hölder inequality,

$$\begin{aligned} w(E) &= |Q| \frac{1}{|Q|} \int_Q w\chi_E \leq |Q| \left( \frac{1}{|Q|} \int_Q w^{r_w} \right)^{\frac{1}{1+\frac{1}{r_w}}} \left( \frac{|E|}{|Q|} \right)^{\frac{1}{r'_w}} \\ &\leq 2w(Q) \left( \frac{|E|}{|Q|} \right)^{\frac{1}{r'_w}} \end{aligned}$$

which yields the desired result, since  $r'_w \simeq c_n[w]_{A_\infty}$ .  $\square$

We recall that the space of bounded mean oscillation functions,  $BMO(\mathbb{R}^n)$ , is the space of locally integrable functions on  $\mathbb{R}^n$ ,  $f$ , such that

$$\|f\|_{BMO} = \sup_Q \frac{1}{|Q|} \int_Q |f(x) - f_Q| dx < \infty$$

where the supremum is taken over all cubes  $Q$  in  $\mathbb{R}^n$  and  $f_Q = \frac{1}{|Q|} \int_Q f(x) dx$ .

Now we recall a fundamental result for us about those functions, the John-Nirenberg theorem.

**Theorem 14** (John-Nirenberg). *For all  $f \in BMO(\mathbb{R}^n)$ , for all cubes  $Q$ , and all  $\alpha > 0$  we have*

$$|\{x \in Q : |f(x) - f_Q| > \alpha\}| \leq e|Q|e^{-\frac{\alpha}{2n\epsilon\|f\|_{BMO}}}.$$

Combining John-Nirenberg Theorem and Lemma 5 we obtain the following result that will be fundamental for our purposes [25].

**Lemma 6.** *Let  $b \in BMO$  and  $w \in A_\infty$ . Then we have that*

$$\|b - b_Q\|_{\exp L(w), Q} \leq c_n[w]_{A_\infty} \|b\|_{BMO}.$$

Furthermore, if  $j > 0$  then

$$\|b - b_Q\|^j_{\exp L^{\frac{1}{j}}(w), Q} \leq c_{n,j}[w]^j_{A_\infty} \|b\|^j_{BMO}.$$

*Proof.* We recall that

$$\|f\|_{\exp L(w),Q} = \inf \left\{ \lambda > 0 : \frac{1}{w(Q)} \int_Q \exp \left( \frac{|f(x)|}{\lambda} \right) - 1 dw < 1 \right\}$$

So it suffices to prove that

$$\frac{1}{w(Q)} \int_Q \exp \left( \frac{|b(x) - b_Q|}{c_n[w]_{A_\infty} \|b\|_{BMO}} \right) - 1 dw < 1,$$

for some  $c_n$  independent of  $w$ ,  $b$  and  $Q$ . Using layer cake formula, Lemma 4 and Theorem 14

$$\begin{aligned} \frac{1}{w(Q)} \int_Q \exp \left( \frac{|f(x)|}{\lambda} \right) - 1 dw &= \frac{1}{w(Q)} \int_0^\infty e^t w(\{x \in Q : |b(x) - b_Q| > \lambda t\}) dt \\ &\leq 2 \frac{1}{w(Q)} \int_0^\infty e^t \left( \frac{|\{x \in Q : |b(x) - b_Q| > \lambda t\}|}{|Q|} \right)^{\frac{1}{c_n[w]_{A_\infty}}} w(Q) dt \\ &\leq 2e \int_0^\infty e^t e^{-\frac{t\lambda}{c_n[w]_{A_\infty} \|b\|_{BMO} e^{2^n}}} dt \end{aligned}$$

So choosing  $\lambda = \alpha c_n e^{2^n} \|b\|_{BMO} [w]_{A_\infty}$

$$2e \int_0^\infty e^t e^{-\frac{t\lambda}{c_n[w]_{A_\infty} \|b\|_{BMO} e^{2^n}}} dt = 2e \int_0^\infty e^{t(1-\alpha)} dt$$

and choosing  $\alpha$  such that the right hand side of the identity is smaller than 1 we are done.

To end the proof of the Lemma we observe that for every measure

$$\frac{1}{\mu(Q)} \int_Q \exp \left( \frac{|f(x)|^j}{\lambda} \right)^{\frac{1}{j}} - 1 d\mu = \frac{1}{\mu(Q)} \int_Q \exp \left( \frac{|f(x)|}{\lambda^{\frac{1}{j}}} \right) - 1 d\mu.$$

Consequently

$$\| |b - b_Q|^j \|_{\exp L^{\frac{1}{j}}(\mu),Q} = \|b - b_Q\|_{\exp L(\mu),Q}^j \quad (3.6)$$

and the desired result follows.  $\square$

#### 4. PROOF OF THE SPARSE DOMINATION

The proof of Theorem 1 follows the scheme in [21] and [24]. We start recalling some basic definitions. Given  $T$  be a sublinear operator we define the grand maximal truncated operator  $\mathcal{M}_T$  by

$$\mathcal{M}_T f(x) = \sup_{Q \ni x} \operatorname{ess\,sup}_{\xi \in Q} |T(f\chi_{\mathbb{R}^n \setminus 3Q})(\xi)|$$

where the supremum is taken over all the cubes  $Q \subset \mathbb{R}^n$  containing  $x$ . We also consider a local version of this operator

$$\mathcal{M}_{T,Q_0} f(x) = \sup_{x \in Q \subseteq Q_0} \operatorname{ess\,sup}_{\xi \in Q} |T(f\chi_{3Q_0 \setminus 3Q})(\xi)|$$

We will need two technical lemmas to prove Theorem 1. The first one is partly a generalization of [21, Lemma 3.2].

**Lemma 7.** *Let  $A$  a Young function such that  $A \in \mathcal{Y}(p_0, p_1)$  with complementary function  $\overline{A}$ . Let  $T$  be an  $\overline{A}$ -Hörmander operator. The following estimates hold*

(1) *For a.e.  $x \in Q_0$*

$$|T(f\chi_{3Q_0})(x)| \leq c_n \|T\|_{L^1 \rightarrow L^{1,\infty}} f(x) + \mathcal{M}_{T,Q_0} f(x) + \varepsilon.$$

(2) For all  $x \in \mathbb{R}^n$

$$\mathcal{M}_T f(x) \leq c_{n,\delta} (H_{\overline{A}} M_A f(x) + M_\delta(Tf)(x) + \|T\|_{L^1 \rightarrow L^{1,\infty}} Mf(x))$$

Furthermore

$$|\{x \in \mathbb{R}^n : \mathcal{M}_T f(x) > \lambda\}| \leq \int_{\mathbb{R}^n} A \left( \max\{c_{A,p_0}, c_{A,p_1}\} c_{n,p_0,p_1} (H_{\overline{A}} + \|T\|_{L^2 \rightarrow L^2}) \frac{|f(x)|}{\lambda} \right) dx. \quad (4.1)$$

*Proof.* (1) was established in [21, Lemma 3.2], so we only have to prove part (2). We are going to follow ideas in [26]. Let  $x, x', \xi \in Q \subset \frac{1}{2} \cdot 3Q$ . Then

$$\begin{aligned} |T(f\chi_{\mathbb{R}^n \setminus 3Q})(\xi)| &= |T(f\chi_{\mathbb{R}^n})(\xi) - T(f\chi_{3Q})(\xi)| \\ &\leq \left| \int_{\mathbb{R}^n \setminus 3Q} (K(\xi, y) - K(x', y)) f(y) dy \right| + |Tf(x')| + |T(f\chi_{3Q})(x')|. \end{aligned}$$

Now we observe that

$$\begin{aligned} &\left| \int_{\mathbb{R}^n \setminus 3Q} (K(\xi, y) - K(x', y)) f(y) dy \right| \\ &\leq \sum_{k=1}^{\infty} 2^{kn} 3^n l(Q)^n \frac{1}{|2^k 3Q|} \int_{2^k 3Q \setminus 2^{k-1} 3Q} |(K(\xi, y) - K(x', y)) f(y)| dy \\ &\leq 2 \sum_{k=1}^{\infty} 2^{kn} 3^n l(Q)^n \left\| (K(\xi, \cdot) - K(x', \cdot)) \chi_{2^k 3Q \setminus 2^{k-1} 3Q} \right\|_{\overline{A}, 2^k 3Q} \|f\|_{A, 2^k 3Q} \\ &\leq c_n H_{\overline{A}} M_A f(x) \end{aligned}$$

Then we have that

$$|T(f\chi_{\mathbb{R}^n \setminus 3Q})(\xi)| \leq c_n H_{\overline{A}} M_A f(x) + |Tf(x')| + |T(f\chi_{3Q})(x')|.$$

$L^\delta$  averaging with  $\delta \in (0, 1)$

$$\begin{aligned} |T(f\chi_{\mathbb{R}^n \setminus 3Q})(\xi)| &\leq c_{n,\delta} \left( H_{\overline{A}} M_A f(x) + \left( \frac{1}{|Q|} \int_Q |Tf(x')|^\delta dx' \right)^{\frac{1}{\delta}} + \left( \frac{1}{|Q|} \int_Q |Tf\chi_{3Q}(x')|^\delta dx' \right)^{\frac{1}{\delta}} \right) \\ &\leq c_{n,\delta} \left( H_{\overline{A}} M_A f(x) + M_\delta(Tf)(x) + \left( \frac{1}{|Q|} \int_Q |Tf\chi_{3Q}(x')|^\delta dx' \right)^{\frac{1}{\delta}} \right). \end{aligned}$$

For the last term we observe that by Kolmogorov's inequality (Lemma 9)

$$\left( \frac{1}{|Q|} \int_Q |Tf\chi_{3Q}(x')|^\delta dx' \right)^{\frac{1}{\delta}} \leq 2 \left( \frac{\delta}{1-\delta} \right)^{\frac{1}{\delta}} \|T\|_{L^1 \rightarrow L^{1,\infty}} \frac{1}{|Q|} \int_{3Q} f \leq c_n \left( \frac{\delta}{1-\delta} \right)^{\frac{1}{\delta}} \|T\|_{L^1 \rightarrow L^{1,\infty}} Mf(x).$$

Summarizing

$$|T(f\chi_{\mathbb{R}^n \setminus 3Q})(\xi)| \leq c_{n,\delta} (H_{\overline{A}} M_A f(x) + M_\delta(Tf)(x) + \|T\|_{L^1 \rightarrow L^{1,\infty}} Mf(x)),$$

and this yields

$$\mathcal{M}_T f(x) \leq c_{n,\delta} (H_{\overline{A}} M_A f(x) + M_\delta(Tf)(x) + \|T\|_{L^1 \rightarrow L^{1,\infty}} Mf(x)). \quad (4.2)$$

Now we observe that  $\|T\|_{L^1 \rightarrow L^{1,\infty}} Mf(x) \leq \|T\|_{L^1 \rightarrow L^{1,\infty}} M_A f(x)$ , and since Lemma 11 provides the following estimate

$$\|T\|_{L^1 \rightarrow L^{1,\infty}} \leq c_n (H_{\overline{A}} + \|T\|_{L^2 \rightarrow L^2}),$$

we have that

$$|\{x \in \mathbb{R}^n : H_{\overline{A}} M_A f(x) + \|T\|_{L^1 \rightarrow L^{1,\infty}} Mf(x) > \lambda\}| \leq c_n \int_{\mathbb{R}^n} A \left( \frac{c_n (H_{\overline{A}} + \|T\|_{L^2 \rightarrow L^2}) |f(x)|}{\lambda} \right) dx. \quad (4.3)$$

Let us focus now on the remaining term. Since  $A \in \mathcal{Y}(p_0, p_1)$  taking into account Lemma 10

$$|\{x \in \mathbb{R}^n : M_\delta(Tf)(x) > \lambda\}| \leq 2 \max\{c_{A,p_0}, c_{A,p_1}\} \int_{\mathbb{R}^n} A\left(\kappa \frac{|f(x)|}{\lambda}\right) dx$$

where  $\kappa = 2 \max\{\|M_\delta \circ T\|_{L^{p_0} \rightarrow L^{p_0,\infty}}, \|M_\delta \circ T\|_{L^{p_1} \rightarrow L^{p_1,\infty}}\}$ . Now we observe that for every  $1 \leq p < \infty$

$$\begin{aligned} \|M_\delta(Tf)\|_{L^{p,\infty}} &= \left\| M(|Tf|^\delta) \right\|_{L^{\frac{p}{\delta},\infty}}^{\frac{1}{\delta}} \leq c_{n,p,\delta} \left\| |Tf|^\delta \right\|_{L^{\frac{p}{\delta},\infty}}^{\frac{1}{\delta}} \\ &= \|Tf\|_{L^{p,\infty}} \leq c_{n,p,\delta} \|T\|_{L^p \rightarrow L^{p,\infty}} \|f\|_{L^p}. \end{aligned}$$

This estimate combined with Lemma 11 yields

$$\|M_\delta \circ T\|_{L^p \rightarrow L^{p,\infty}} \leq c_{n,p,\delta} (H_{\overline{A}} + \|T\|_{L^2 \rightarrow L^2}).$$

Hence

$$|\{x \in \mathbb{R}^n : M_\delta(Tf)(x) > \lambda\}| \leq 2 \max\{c_{A,p_0}, c_{A,p_1}\} \int_{\mathbb{R}^n} A\left(c_{n,p_0,p_1,\delta} (H_{\overline{A}} + \|T\|_{L^2 \rightarrow L^2}) \frac{|f(x)|}{\lambda}\right) dx. \quad (4.4)$$

Since  $\frac{A(t)}{t}$  is non decreasing, it is not hard to see that for  $c \geq 1$   $cA(t) \leq A(ct)$ . Using this fact combined with equations (4.2), (4.3) and (4.4) we obtain (4.1).  $\square$

*Proof of Theorem 1.* From Remark 1 it follows that there exist  $3^n$  dyadic lattices such that for every cube  $Q$  of  $\mathbb{R}^n$  there is a cube  $R_Q \in \mathcal{D}_j$  for some  $j$  for which  $3Q \subset R_Q$  and  $|R_Q| \leq 9^n|Q|$

We fix a cube  $Q_0 \subset \mathbb{R}^n$ . We claim that there exists a  $\frac{1}{2}$ -sparse family  $\mathcal{F} \subseteq \mathcal{D}(Q_0)$  such that for a.e.  $x \in Q_0$

$$|T_b^m(f\chi_{3Q_0})(x)| \leq c_n C_T \sum_{h=0}^m \binom{m}{h} \mathcal{B}_{\mathcal{F}}^{m,h}(b, f)(x) \quad (4.5)$$

where

$$\mathcal{B}_{\mathcal{F}}^{m,h}(b, f)(x) = \sum_{Q \in \mathcal{F}} |b(x) - b_{R_Q}|^{m-h} \|f\|_{A,3Q} |b - b_{R_Q}|^h \chi_Q(x)$$

Suppose that we have already proved (4.5). Let us take a partition of  $\mathbb{R}^n$  by cubes  $Q_j$  such that  $\text{supp}(f) \subseteq 3Q_j$  for each  $j$ . We can do it as follows. We start with a cube  $Q_0$  such that  $\text{supp}(f) \subset Q_0$ . And cover  $3Q_0 \setminus Q_0$  by  $3^n - 1$  congruent cubes  $Q_j$ . Each of them satisfies  $Q_0 \subset 3Q_j$ . We do the same for  $9Q_0 \setminus 3Q_0$  and so on. The union of all those cubes, including  $Q_0$ , will satisfy the desired properties.

We apply the claim to each cube  $Q_j$ . Then we have that since  $\text{supp } f \subseteq 3Q_j$  the following estimate holds a.e.  $x \in Q_j$

$$|T_b^m f(x)| \chi_{Q_j}(x) = |T_b^m(f\chi_{3Q_j})(x)| \leq c_n C_T \mathcal{B}_{\mathcal{F}_j}^{m,h}(b, f)(x)$$

where each  $\mathcal{F}_j \subseteq \mathcal{D}(Q_j)$  is a  $\frac{1}{2}$ -sparse family. Taking  $\mathcal{F} = \bigcup \mathcal{F}_j$  we have that  $\mathcal{F}$  is a  $\frac{1}{2}$ -sparse family and

$$|T_b^m f(x)| \leq c_n C_T \sum_{h=0}^m \binom{m}{h} \mathcal{B}_{\mathcal{F}}^{m,h}(b, f)(x)$$

Now since  $3Q \subset R_Q$  and  $|R_Q| \leq 3^n|3Q|$  we have that  $\|f\|_{A,3Q} \leq c_n \|f\|_{A,R_Q}$ . Setting

$$\mathcal{S}_j = \{R_Q \in \mathcal{D}_j : Q \in \mathcal{F}\}$$

and using that  $\mathcal{F}$  is  $\frac{1}{2}$ -sparse, we obtain that each family  $\mathcal{S}_j$  is  $\frac{1}{2 \cdot 9^n}$ -sparse. Then we have that

$$|T_b^m f(x)| \leq c_{n,m} C_T \sum_{j=1}^{3^n} \sum_{h=0}^m \binom{m}{h} \mathcal{A}_{\mathcal{S}_j}^{m,h}(b, f)(x)$$

*Proof of the claim (4.5).* To prove the claim it suffices to prove the following recursive estimate: There exist pairwise disjoint cubes  $P_j \in \mathcal{D}(Q_0)$  such that  $\sum_j |P_j| \leq \frac{1}{2}|Q_0|$  and

$$\begin{aligned} & |T_b^m(f\chi_{3Q_0})(x)|\chi_{Q_0} \\ & \leq c_n C_T \sum_{h=0}^m \binom{m}{h} |b(x) - b_{R_{Q_0}}|^{m-h} \|f(b - b_{R_{Q_0}})^h\|_{3Q_0} \chi_{Q_0}(x) \\ & + \sum_j |T_b^m(f\chi_{3P_j})(x)|\chi_{P_j}. \end{aligned}$$

a.e. in  $Q_0$ . Iterating this estimate we obtain (4.5) with  $\mathcal{F} = \{P_j^k\}$  where  $\{P_j^0\} = \{Q_0\}$ ,  $\{P_j^1\} = \{P_j\}$  and  $\{P_j^k\}$  are the cubes obtained at the  $k$ -th stage of the iterative process. It is also clear that  $\mathcal{F}$  is a  $\frac{1}{2}$ -sparse family. Indeed, for each  $P_j^k$  it suffices to choose

$$E_{P_j^k} = P_j^k \setminus \bigcup_j P_j^{k+1}.$$

Let us prove then the recursive estimate. We observe that for any arbitrary family of disjoint cubes  $P_j \in \mathcal{D}(Q_0)$  we have that

$$\begin{aligned} & |T_b^m(f\chi_{3Q_0})(x)|\chi_{Q_0}(x) \\ & \leq |T_b^m(f\chi_{3Q_0})(x)|\chi_{Q_0 \setminus \bigcup_j P_j}(x) + \sum_j |T_b^m(f\chi_{3Q_0})(x)|\chi_{P_j}(x) \\ & \leq |T_b^m(f\chi_{3Q_0})(x)|\chi_{Q_0 \setminus \bigcup_j P_j}(x) + \sum_j \left| T_b^m(f\chi_{3Q_0 \setminus 3P_j})(x) \right| \chi_{P_j}(x) + \sum_j |T_b^m(f\chi_{3P_j})(x)|\chi_{P_j}(x) \end{aligned}$$

So it suffices to show that we can choose a family of pairwise disjoint cubes  $P_j \in \mathcal{D}(Q_0)$  with  $\sum_j |P_j| \leq \frac{1}{2}|Q_0|$  and such that for a.e.  $x \in Q_0$

$$\begin{aligned} & |T_b^m(f\chi_{3Q_0})(x)|\chi_{Q_0 \setminus \bigcup_j P_j}(x) + \sum_j \left| T_b^m(f\chi_{3Q_0 \setminus 3P_j})(x) \right| \chi_{P_j}(x) \\ & \leq c_n C_T \sum_{h=0}^m \binom{m}{h} |b(x) - b_{R_{Q_0}}|^{m-h} \|f(b - b_{R_{Q_0}})^h\|_{3Q} \chi_Q(x) \end{aligned}$$

Using that  $T_b^m f = T_{b-c}^m f$  for any  $c \in \mathbb{R}$ , and also that

$$T_{b-c}^m f = \sum_{h=0}^m (-1)^h \binom{m}{h} T((b-c)^h f)(b-c)^{m-h}$$

we obtain

$$\begin{aligned} & |T_b^m(f\chi_{3Q_0})|\chi_{Q_0 \setminus \bigcup_j P_j} + \sum_j |T_b^m(f\chi_{3Q_0 \setminus 3P_j})|\chi_{P_j} \\ & \leq \sum_{h=0}^m \binom{m}{h} |b - b_{R_{Q_0}}|^{m-h} |T((b - b_{R_{Q_0}})^h f \chi_{3Q_0})|\chi_{Q_0 \setminus \bigcup_j P_j} \end{aligned} \quad (4.6)$$

$$+ \sum_{h=0}^m \binom{m}{h} |b - b_{R_{Q_0}}|^{m-h} \sum_j |T((b - b_{R_{Q_0}})^h f \chi_{3Q_0 \setminus 3P_j})|\chi_{P_j}. \quad (4.7)$$



Now for  $h = 0, 1, \dots, m$  we define the set  $E_h$  as

$$E_h = \left\{ x \in Q_0 : |b - b_{R_{Q_0}}|^h |f| > \alpha_n \|b - b_{R_{Q_0}}\|^h f\|_{A,3Q_0} \right\} \\ \cup \left\{ x \in Q_0 : \mathcal{M}_{T,Q_0} \left( |b - b_{R_{Q_0}}|^h f \right) > \alpha_n C_T \|b - b_{R_{Q_0}}\|^h f\|_{A,3Q_0} \right\}$$

and we call  $E = \bigcup_{h=0}^m E_h$ . Now we note that taking into account the convexity of  $A$  and the second part in Lemma 7,

$$|E_h| \leq \frac{\int_{Q_0} |b - b_{R_{Q_0}}|^h |f|}{\alpha_n \|f\|_{A,3Q_0}} + c_n \int_{3Q_0} A \left( \frac{\max\{c_{A,p_0}, c_{A,p_1}\} c_{n,p_0,p_1} (H_{\overline{A}} + \|T\|_{L^2 \rightarrow L^2}) |b - b_{R_{Q_0}}|^h |f|}{\alpha_n C_T \|b - b_{R_{Q_0}}\|^h f\|_{A,3Q_0}} \right) dx \\ \leq 3^n \frac{1}{|3Q_0|} \frac{\int_{3Q_0} |b - b_{R_{Q_0}}|^h |f|}{\alpha_n \|b - b_{R_{Q_0}}\|^h f\|_{A,3Q_0}} |Q_0| + \frac{c_n}{\alpha_n} |Q_0| \frac{1}{|3Q_0|} \int_{3Q_0} A \left( \frac{|b - b_{R_{Q_0}}|^h |f|}{\|b - b_{R_{Q_0}}\|^h f\|_{A,3Q_0}} \right) dx \\ \leq \left( \frac{2 \cdot 3^n}{\alpha_n} + \frac{c_n}{\alpha_n} \right) |Q_0|.$$

Then, choosing  $\alpha_n$  big enough, we have that

$$|E| \leq \frac{1}{2^{n+2}} |Q_0|.$$

Now we apply Calderón-Zygmund decomposition to the function  $\chi_E$  on  $Q_0$  at height  $\lambda = \frac{1}{2^{n+1}}$ . We obtain pairwise disjoint cubes  $P_j \in \mathcal{D}(Q_0)$  such that

$$\chi_E(x) \leq \frac{1}{2^{n+1}}$$

for a.e.  $x \notin \bigcup_j P_j$ . From this it follows that  $|E \setminus \bigcup_j P_j| = 0$ . And also that family satisfies that

$$\sum_j |P_j| = \left| \bigcup_j P_j \right| \leq 2^{n+1} |E| \leq \frac{1}{2} |Q_0|$$

and also that

$$\frac{1}{2^{n+1}} \leq \frac{1}{|P_j|} \int_{P_j} \chi_E(x) = \frac{|P_j \cap E|}{|P_j|} \leq \frac{1}{2}$$

from which it readily follows that  $|P_j \cap E^c| > 0$ .

We observe that then for each  $P_j$  we have that since  $P_j \cap E^c \neq \emptyset$ ,  $\mathcal{M}_{T,Q_0} \left( |b - b_{R_{Q_0}}|^h f \right) (x) \leq \alpha_n C_T \|b - b_{R_{Q_0}}\|^h f\|_{A,3Q_0}$  for some  $x \in P_j$  and this implies

$$\operatorname{ess\,sup}_{\xi \in Q} \left| T(|b - b_{R_{Q_0}}|^h f \chi_{3Q_0 \setminus 3Q})(\xi) \right| \leq \alpha_n C_T \|b - b_{R_{Q_0}}\|^h f\|_{A,3Q_0}$$

which allows us to control the summation in (4.7).

Now, by (1) in Lemma (7) since by Lemma 11  $\|T\|_{L^1 \rightarrow L^{1,\infty}} \leq c_n (H_A + \|T\|_{L^2 \rightarrow L^2})$  we know that a.e.  $x \in Q_0$ ,

$$\left| T(|b - b_{R_{Q_0}}|^h |f| \chi_{3Q_0})(x) \right| \leq c_n C_T |b(x) - b_{R_{Q_0}}|^h |f(x)| + \mathcal{M}_{T,Q_0} \left( |b - b_{R_{Q_0}}|^h |f| \right) (x)$$

Since  $|E \setminus \bigcup_j P_j| = 0$ , we have that, by the definition of  $E$ , the following estimate

$$|b(x) - b_{R_{Q_0}}|^h |f(x)| \leq \alpha_n \|b - b_{R_{Q_0}}\|^h f\|_{A,3Q_0}$$

holds a.e.  $x \in Q_0 \setminus \bigcup_j P_j$  and also

$$\mathcal{M}_{T,Q_0} \left( |b - b_{R_{Q_0}}|^h |f| \right) (x) \leq \alpha_n \|b - b_{R_{Q_0}}\|^h f\|_{A,3Q_0}$$

holds a.e.  $x \in Q_0 \setminus \bigcup_j P_j$ . Consequently

$$\left| T((b - b_{R_{Q_0}})^h f \chi_{3Q_0})(x) \right| \leq c_n C_T \| |b - b_{R_{Q_0}}|^h f \|_{A, 3Q_0}.$$

Those estimates allow us to control the remaining terms in (4.6) so we are done.

## 5. PROOFS OF STRONG TYPE ESTIMATES

**5.1. Proof of Theorem 2.** We establish first the corresponding estimate for  $T$ . Combining [2, Lemma 4.1] with [15, Theorem 1.1] and taking into account the sparse domination

$$\begin{aligned} \|Tf\|_{L^p(w)} &\leq c_n c_T \sum_{j=1}^{3^n} \left( \int_{\mathbb{R}^n} (\mathcal{A}_{A, S_j} f)^p \right)^{1/p} = c_n c_T \sum_{j=1}^{3^n} \left( \int_{\mathbb{R}^n} \left( \sum_{Q \in \mathcal{S}} \|f\|_{A, Q} \chi_Q(x) \right)^p dx \right)^{1/p} \\ &\leq c_n c_T \sum_{j=1}^{3^n} \mathcal{K}_{r, A} \left( \int_{\mathbb{R}^n} \left( \sum_{Q \in \mathcal{S}} \left( \frac{1}{|Q|} \int_Q |f|^r \right)^{1/r} \chi_Q(x) \right)^p dx \right)^{1/p} \\ &= c_n c_T \sum_{j=1}^{3^n} \mathcal{K}_{r, A} \|\mathcal{A}_S^{1/r}(|f|^r)\|_{L^{p/r}(w)}^{1/r} \\ &\leq c_n c_T \mathcal{K}_{r, A} [w]_{A_{p/r}}^{\frac{1}{p/r} \frac{1}{r}} \left( [w]_{A_\infty}^{\left(r - \frac{r}{p}\right) \frac{1}{r}} + [\sigma]_{A_\infty}^{\frac{1}{p/r} \frac{1}{r}} \right) \| |f|^r \|_{L^{p/r}(w)}^{1/r} \\ &= c_n c_T \mathcal{K}_{r, A} [w]_{A_{p/r}}^{\frac{1}{p}} \left( [w]_{A_\infty}^{\frac{1}{p'}} + [\sigma]_{A_\infty}^{\frac{1}{p}} \right) \|f\|_{L^p(w)} \end{aligned}$$

Now for the commutator and the iterated commutator we use the conjugation method (See [4, 3, 39] for more details about this method). We recall that

$$T_b^m f = \frac{m!}{2\pi i} \int_{|z|=\varepsilon} \frac{e^{bz} T(e^{-bz} f)}{z^{m+1}} dz.$$

If  $w \in A_{p/r}$ , taking norms

$$\begin{aligned} \|T_b^m f\|_{L^p(w)} &\leq \frac{m!}{2\pi \varepsilon^m} \sup_{|z|=\varepsilon} \|e^{bz} T(f e^{-bz})\|_{L^p(w)} \\ &= \frac{m!}{2\pi \varepsilon^m} \sup_{|z|=\varepsilon} \|T(f e^{-bz})\|_{L^p(e^{\operatorname{Re}(bz)p} w)} \\ &\leq c_n c_T \mathcal{K}_{r, A} \frac{m!}{2\pi \varepsilon^m} \sup_{|z|=\varepsilon} [e^{\operatorname{Re}(bz)p} w]_{A_{p/r}}^{\frac{1}{p}} \left( [e^{\operatorname{Re}(bz)p} w]_{A_\infty}^{\frac{1}{p'}} + [e^{-\operatorname{Re}(bz) \frac{p}{p/r-1}} \sigma]_{A_\infty}^{\frac{1}{p}} \right) \|f\|_{L^p(w)}. \end{aligned}$$

Now taking into account [14, Lemma 2.1] and [16, Lemma 7.3] we have that  $[e^{\operatorname{Re}(bz)p} w]_{A_{p/r}} \leq c_{n, p/r} [w]_{A_{p/r}}$ ,  $[e^{\operatorname{Re}(bz)p} w]_{A_\infty} \leq c_n [w]_{A_\infty}$  and  $[e^{-\operatorname{Re}(bz) \frac{p}{p/r-1}} \sigma]_{A_\infty} \leq c_n [\sigma]_{A_\infty}$  provided that

$$|z| \leq \frac{\varepsilon_{n, p}}{\|b\|_{BMO} ([w]_{A_\infty} + [\sigma]_{A_\infty})}.$$

This yields

$$\|T_b^m f\|_{L^p(w)} \leq c_{n, m} c_T \mathcal{K}_{r, A} [w]_{A_{p/r}}^{\frac{1}{p}} \left( [w]_{A_\infty}^{\frac{1}{p'}} + [\sigma_{p/r}]_{A_\infty}^{\frac{1}{p}} \right) ([w]_{A_\infty} + [\sigma_{p/r}]_{A_\infty})^m \|b\|_{BMO}^m \|f\|_{L^p(w)}.$$

5.2. **Proof of Theorem 3.** Using duality we have that

$$\|A_{B,S}f\|_{L^p(w)} = \sup_{\|g\|_{L^{p'}(w)}=1} \sum_{Q \in S} \|f\|_{B,Q} \int_Q gw.$$

Now we observe that, since  $\frac{1}{w(Q)} \int_Q gw \leq \inf_Q M_w^d(g)$ , then

$$\begin{aligned} \sum_{Q \in S} \left( \frac{1}{w(Q)} \int_Q g \right)^{p'} w(E_Q) &\leq \sum_{Q \in S} \int_{E_Q} M_w^d(g)^{p'} w \\ &\leq \int_{\mathbb{R}^n} M_w^d(g)^{p'} w \leq c_{n,p} \|g\|_{L^{p'}(w)}^{p'} \end{aligned} \quad (5.1)$$

Since we know that  $A^{-1}(t)\overline{B}^{-1}(t)C^{-1}(t) \leq ct$  for every  $t \geq t_0$ , some  $t_0 > 0$ , applying generalized Hölder inequality we have that

$$\|f\|_{B,Q} = \|fw^{\frac{1}{p}}w^{-\frac{1}{p}}\|_{B,Q} \leq \tilde{c}_1 \|fw^{\frac{1}{p}}\|_{A,Q} \|w^{-\frac{1}{p}}\|_{C,Q}$$

Since  $A \in B_p$ , we have

$$\begin{aligned} \sum_{Q \in S} \|fw^{\frac{1}{p}}\|_{A,Q}^p |E_Q| &\leq \sum_{Q \in S} \int_{E_Q} M_A(fw^{\frac{1}{p}})^p \\ &\leq \int_{\mathbb{R}^n} M_A(fw^{\frac{1}{p}})^p \\ &\leq c_{n,p} \int_{\mathbb{R}^n} (fw^{\frac{1}{p}})^p = c_{n,p} \|f\|_{L^p(w)}^p. \end{aligned} \quad (5.2)$$

Then, taking into account (5.2) and (5.1),

$$\begin{aligned} \sum_{Q \in S} \|f\|_{B,Q} \int_Q g &\leq \sum_{Q \in S} \|fw^{\frac{1}{p}}\|_{A,Q} \|w^{-\frac{1}{p}}\|_{C,Q} \int_Q g \\ &= c_{n,p} \sum_{Q \in S} \|fw^{\frac{1}{p}}\|_{A,Q} |E_Q|^{\frac{1}{p}} \frac{\|w^{-\frac{1}{p}}\|_{C,Q}}{|E_Q|^{\frac{1}{p}}} \frac{w(Q)}{w(E_Q)^{\frac{1}{p'}}} \left( \frac{1}{w(Q)} \int_Q g \right) w(E_Q)^{\frac{1}{p'}} \\ &\leq c_{n,p} \sup_Q \left\{ \frac{\|w^{-\frac{1}{p}}\|_{C,Q}}{|E_Q|^{\frac{1}{p}}} \frac{w(Q)}{w(E_Q)^{\frac{1}{p'}}} \right\} \left( \sum_{Q \in S} \|fw^{\frac{1}{p}}\|_{A,Q} |E_Q| \right)^{\frac{1}{p}} \left[ \left( \sum_{Q \in S} \left( \frac{1}{w(Q)} \int_Q g \right)^{p'} w(E_Q) \right) \right]^{\frac{1}{p'}} \\ &\leq c_{n,p} \sup_Q \left\{ \frac{\|w^{-\frac{1}{p}}\|_{C,Q}}{|E_Q|^{\frac{1}{p}}} \frac{w(Q)}{w(E_Q)^{\frac{1}{p'}}} \right\} \|f\|_{L^p(w)} \|g\|_{L^{p'}(w)}, \end{aligned}$$

and we are left with controlling the supremum. To do that we observe that taking into account that

$$w(Q) \leq c[w]_{A_p} w(E_Q)$$

we have that

$$\begin{aligned}
\frac{\|w^{-\frac{1}{p}}\|_{C,Q}}{|E_Q|^{\frac{1}{p}}} \frac{w(Q)}{w(E_Q)^{\frac{1}{p'}}} &= \|w^{-1/p}\|_{C,Q} \frac{w(Q)^{1/p}}{|E_Q|^{1/p}} \frac{w(Q)^{1/p'}}{w(E_Q)^{1/p'}} \\
&= c_p \|w^{-\frac{1}{p}}\|_{C,Q} \frac{w(Q)^{1/p}}{|Q|^{1/p}} \frac{w(Q)^{1/p'}}{w(E_Q)^{1/p'}} \\
&\leq c_p [w]_{A_p(C)}^{\frac{1}{p}} \frac{w(Q)^{1/p'}}{w(E_Q)^{1/p'}} \\
&\leq c_{n,p,\eta} [w]_{A_p(C)}^{\frac{1}{p}} [w]_{A_p}^{\frac{1}{p'}}.
\end{aligned}$$

This ends the proof of the Theorem.

## 6. PROOFS OF COIFMAN-FEFFERMAN ESTIMATES AND RELATED RESULTS

**6.1. Proof of Theorem 4.** The case  $m = 0$  is easier than the case  $m > 0$ . It suffices to repeat the same proof that we provide here for the case  $m > 0$  with the obvious modifications so we omit it. Let then  $m > 0$ . Using Theorem 1 it suffices to control each  $\mathcal{A}_{A,S}^{m,h}(b, f)$ . We observe that taking into account Lemma 6 and Hölder inequality,

$$\begin{aligned}
\int_{\mathbb{R}^n} \mathcal{A}_{B,S}^{m,h}(b, f) g w dx &= \sum_{Q \in \mathcal{S}} \frac{1}{w(Q)} \int_Q |b(x) - b_Q|^{m-h} g(x) w(x) dx w(Q) \| (b - b_Q)^h f \|_{B,Q} \\
&\leq \sum_{Q \in \mathcal{S}} \| (b - b_Q)^{m-h} \|_{\exp L^{\frac{1}{m-h}}(w), Q} \| g \|_{L(\log L)^{m-h}(w), Q} w(Q) \| (b - b_Q)^h \|_{\exp L^{\frac{1}{h}}, Q} \| f \|_{A,Q} \\
&\leq c_n [w]_{A_\infty}^{m-h} \| b \|_{\text{BMO}}^m \sum_{Q \in \mathcal{S}} \| g \|_{L(\log L)^{m-h}(w), Q} \| f \|_{A,Q} w(Q)
\end{aligned}$$

Now we observe that

$$\begin{aligned}
&\sum_{Q \in \mathcal{S}} \| g \|_{L(\log L)^{m-h}(w), Q} \| f \|_{A,Q} w(Q) \\
&\leq \sum_{F \in \mathcal{F}} \| g \|_{L(\log L)^{m-h}(w), F} \| f \|_{A,F} \sum_{Q \in \mathcal{S}, \pi(Q)=F} w(Q) \\
&\leq c_n [w]_{A_\infty} \sum_{F \in \mathcal{F}} \| g \|_{L(\log L)^{m-h}(w), F} \| f \|_{A,F} w(F) \\
&\leq c_n [w]_{A_\infty} \int_{\mathbb{R}^n} (M_A f) (M_{L \log L^{m-h}(w)} g) w dx \\
&\leq c_n [w]_{A_\infty} \int_{\mathbb{R}^n} (M_A f) (M_w^{m-h+1} g) w dx
\end{aligned}$$

where  $\mathcal{F}$  is the family of the principal cubes in the usual sense, namely,

$$\mathcal{F} = \cup_{k=0}^{\infty} \mathcal{F}_k$$

with  $\mathcal{F}_0 := \{\text{maximal cubes in } \mathcal{S}\}$  and

$$\mathcal{F}_{k+1} := \bigcup_{F \in \mathcal{F}_k} \text{ch}_{\mathcal{F}}(F), \quad \text{ch}_{\mathcal{F}}(F) = \{Q \subsetneq F \text{ maximal s.t. } \tau(Q) > 2\tau(F)\}$$

where  $\tau(Q) = \|g\|_{L(\log L)^{m-h}(w), Q} \|f\|_{A,Q}$  and  $\pi(Q)$  is the minimal principal cube which contains  $Q$ .

For the case  $p = 1$  we take  $g = 1$ . Then

$$\int_{\mathbb{R}^n} (M_A f) (M_w^{m-h+1} g) w dx \leq \int_{\mathbb{R}^n} (M_A f) w dx$$

and we are done. In the case  $p > 1$  we have that

$$\begin{aligned} \int_{\mathbb{R}^n} (M_A f)(M_w^{m-h+1} g) w dx &\leq \|M_A f\|_{L^p(w)} \|M_w^{m-h+1} g\|_{L^{p'}(w)} \\ &\leq c_n p^{m-h+1} \|M_A f\|_{L^p(w)} \|g\|_{L^{p'}(w)} \end{aligned}$$

and taking supremum on  $\|g\|_{L^{p'}(w)} = 1$  we end the proof.

**6.2. Proof of theorem 5.** We are going to follow the scheme of the proof of [31, Theorem 3.2]. We consider the kernel that appears in [30, Theorem 5]

$$k(t) = A^{-1} \left( \frac{1}{t^n (1 - \log t)^{1+\beta}} \right) \chi_{(0,1)}(t).$$

We observe that  $K(x) = k(|x|) \in L^1(\mathbb{R}^n)$ . Indeed

$$\begin{aligned} &A \left( \frac{1}{|B(0,1)|} \int_{\mathbb{R}^n} A^{-1} \left( |x|^{-n} \left( \log \frac{e}{|x|} \right)^{-(1+\beta)} \chi_{(0,1)}(|x|) \right) dx \right) \\ &= A \left( \frac{1}{|B(0,1)|} \int_{|x|<1} A^{-1} \left( |x|^{-n} \left( \log \frac{e}{|x|} \right)^{-(1+\beta)} \right) dx \right) \\ &\leq \frac{1}{|B(0,1)|} \int_{|x|<1} |x|^{-n} \left( \log \frac{e}{|x|} \right)^{-(1+\beta)} dx \leq c_{n,\beta} \end{aligned}$$

Then

$$\int_{\mathbb{R}^n} A^{-1} \left( |x|^{-n} \left( \log \frac{e}{|x|} \right)^{-(1+\beta)} \chi_{(0,1)}(|x|) \right) dx \leq A^{-1}(c_{n,\beta}) |B(0,1)|$$

and hence  $K(x) \in L^1$ . Now we define  $\tilde{K}(x) = K(x - \eta)$  where  $\eta \in \mathbb{R}^n$  is taken away from the origin, let's say  $|\eta| = 4$ . We define

$$Tf(x) = \tilde{K} * f(x) = \int_{\mathbb{R}^n} K(x - \eta - y) f(y) dy.$$

Since  $\tilde{K} \in L^1$  we have that  $T : L^q \rightarrow L^q$  for every  $1 < q < \infty$ . We observe now that the kernel  $\tilde{K}$  satisfies an  $A$ -Hörmander condition [30, Theorem 5].

Let us assume that  $T$  maps  $L^p(\mathbb{R}^n, w)$  into  $L^{p,\infty}(\mathbb{R}^n, w)$ . We define

$$f(x) = |x + \eta|^{-\frac{\gamma_1 n}{p}} \chi_{\{|x+\eta|<1\}}(x) \in L^p(\mathbb{R}^n)$$

with  $\gamma_1 \in (0, 1)$  to be chosen. If  $|x + \eta| < 1$  then  $3 < |x| < 5$  and therefore

$$\begin{aligned} \sup_{\lambda>0} \lambda^p w \{x \in \mathbb{R}^n : |Tf(x)| > \lambda\} &\leq c \left( \int_{\mathbb{R}^n} |f(x)| w(x) dx \right) \\ &\leq c \frac{1}{3^{n\alpha}} \left( \int_{\mathbb{R}^n} |f(x)| dx \right) < \infty \end{aligned} \tag{6.1}$$

Now we consider

$$\tilde{h}(t) = k(t) t^{-\frac{\gamma_1 n}{p} + n} \quad t \in (0, 1).$$

Let us call  $\psi(t) = \frac{1}{t^n(1-\log t)^{1+\beta}}$ . Then we have that  $k(t) = A^{-1}(\psi(t))$ . Let us choose  $0 < s < \min \left\{ \frac{1}{3r'}, \frac{\gamma_1}{p} \right\}$ . Then, since  $\varphi(u) < \kappa_s u^s$  for some  $t > c_s$ , we observe that taking  $t$  small enough we

have that  $\frac{1}{t^n(1-\log t)^{1+\beta}} > \max\{c_0, c_s\}$ . This yields that

$$\begin{aligned}\tilde{h}(t) &= A^{-1} \left( \frac{1}{t^n(1-\log t)^{1+\beta}} \right) t^{-\frac{\gamma_1 n}{p}+n} \simeq \frac{1}{t^{\frac{n}{r}}(1-\log t)^{\frac{1+\beta}{r}} \varphi \left( \frac{1}{t^n(1-\log t)^{1+\beta}} \right)} t^{-\frac{\gamma_1 n}{p}+n} \\ &\geq \frac{1}{\kappa_s} \frac{1}{(1-\log t)^{\frac{1+\beta}{r}} \left( \frac{1}{t^n(1-\log t)^{1+\beta}} \right)^s} t^{-\frac{\gamma_1 n}{p}} = \frac{1}{\kappa_s} (1-\log t)^{(1+\beta)(s-\frac{1}{r})} t^{-\frac{\gamma_1 n}{p}+ns} = \frac{1}{\kappa_s} h(t)\end{aligned}$$

and we have that  $h$  is decreasing for  $t$  small enough that let's say  $t \in (0, t_0)$ . Actually, we choose  $t_0$  such that  $k(t)$  is decreasing in that range as well. Now we take  $\delta_0 = \frac{2}{3}t_0$ . We observe that for  $|x| < \delta_0$ ,

$$\begin{aligned}Tf(x) &= \int_{|\eta+y|<1} K(x-\eta-y)|y+\eta|^{-\frac{\gamma_1 n}{p}} dy \\ &= \int_{|y|<1} K(x-y)|y|^{-\frac{\gamma_1 n}{p}} dy = \int_{|y|<1} k(|x-y|)|y|^{-\frac{\gamma_1 n}{p}} dy \\ &\geq k\left(\frac{3}{2}|x|\right) \int_{|y|<\frac{|x|}{2}} |y|^{-\frac{\gamma_1 n}{p}} dy \geq k\left(\frac{3}{2}|x|\right) \frac{|x|^{-\frac{\gamma_1 n}{p}}}{2^{-\frac{\gamma_1 n}{p}}} |x|^n \\ &\geq c\tilde{h}\left(\frac{3|x|}{2}\right) \geq c\frac{1}{\kappa_s} h\left(\frac{3|x|}{2}\right).\end{aligned}$$

This yields

$$\begin{aligned}\sup_{\lambda>0} \lambda^p w \{x \in \mathbb{R}^n : |Tf(x)| > \lambda\} &\geq \sup_{\lambda>0} \lambda^p w \left\{ |x| < \delta_0 : c\frac{1}{\kappa_s} h\left(\frac{3|x|}{2}\right) > \lambda \right\} \\ &\geq c_s \sup_{\lambda>h(t_0)} \lambda^p w \left\{ |x| < \delta_0 : h\left(\frac{3|x|}{2}\right) > \lambda \right\} \\ &\geq c_s \sup_{0<t<t_0} h(t)^p w \left\{ |x| < \frac{2t}{3} \right\} \\ &= c_s \sup_{0<t<t_0} (1-\log t)^{(1+\beta)(s-\frac{1}{r})p} t^{-\gamma_1 n+pn s} \int_{|y|<\frac{2t}{3}} |x|^{-\gamma_1 n} dy \\ &\simeq c_s \sup_{0<t<t_0} (1-\log t)^{(1+\beta)(\frac{1}{2}-p)} t^{-\gamma_1 n+pn s+n-\gamma_1 n}\end{aligned}$$

Now we observe that

$$-\gamma_1 n + pn s + n - \gamma_1 n < 0 \iff n(-\gamma_1 + ps - \gamma + 1) < 0 \iff 1 + ps < \gamma_1 + \gamma.$$

Since  $s < \frac{1}{3r'}$ , choosing  $\gamma_1 = 1 - \frac{p}{r'2}$  we have that

$$\gamma_1 + \gamma = 1 - \frac{p}{r'2} + \gamma > 1 - \frac{p}{r'2} + \frac{p}{r'} = 1 + \frac{p}{2r'}.$$

Consequently

$$1 + ps < \gamma_1 + \gamma$$

and this yields

$$\sup_{\lambda>0} \lambda^p w \{x \in \mathbb{R}^n : |Tf(x)| > \lambda\} = \infty$$

Which contradicts (6.1).

**6.3. Proof of Theorem 6.** Assume that (1.9) holds for  $M_B$  with  $B(t) \leq ct^q$  for every  $t \geq c$  and  $1 < q < r'$ . Arguing as in [31, Proof of Theorem 3.1], it suffices to disprove the estimate for some  $0 < p_0 < \infty$ . Let us choose  $q < p_0 < r'$ . Assume that for every  $w \in A_1 \subseteq A_\infty$  we have that  $\|Tf\|_{L^{p_0,\infty}(w)} \leq c\|M_B f\|_{L^{p_0,\infty}(w)}$ . Then we observe that

$$\|Tf\|_{L^{p_0,\infty}(w)} \leq c\|M_B f\|_{L^{p_0,\infty}(w)} \leq c\|M_q f\|_{L^{p_0,\infty}(w)} \leq c\|f\|_{L^{p_0,\infty}(w)}.$$

and this in particular holds for the weight  $w(x) = |x|^{-n\gamma}$  with  $\gamma \in (\frac{p_0}{r'}, 1)$  contradicting Theorem 5.

## 7. PROOFS OF ENDPOINT ESTIMATES

The proofs that we present in this section will follow the strategy outlined in [10] and generalized in [24]. Let  $A$  be a Young function satisfying

$$A(4t) \leq \Lambda_A A(t) \quad (t > 0, \Lambda_A \geq 1). \quad (7.1)$$

Let  $\mathcal{D}$  be a dyadic lattice and  $k \in \mathbb{N}$ . We denote

$$\mathcal{F}_k = \left\{ Q \in \mathcal{D} : 4^{k-1} < \|f\|_{A,Q} \leq 4^k \right\}$$

Now we recall [24, Lemma 4.3].

**Lemma 8.** *Suppose that the family  $\mathcal{F}_k$  is  $\left(1 - \frac{1}{2\Lambda_A}\right)$ -sparse. Let  $w$  weight and let  $E$  be an arbitrary measurable set with  $w(E) < \infty$ . Then for every Young function  $\varphi$ ,*

$$\int_E \left( \sum_{Q \in \mathcal{F}_k} \chi_Q \right) w dx \leq 2^k w(E) + \frac{4\Lambda_A}{\bar{\varphi}^{-1}((2\Lambda_A)^{2^k})} \int_{\mathbb{R}^n} A(4^k |f|) M_\varphi w dx.$$

Using the preceding Lemma we are in the position to prove Theorem 7.

**7.1. Proof of Theorem 7.** Let

$$E = \left\{ x \in \mathbb{R}^n : \mathcal{A}_{\mathcal{S},A} f(x) > 4, M_A f(x) \leq \frac{1}{4} \right\}.$$

By homogeneity, taking into account Lemma 3, it suffices to prove that

$$w(E) \leq c\kappa_\varphi \int_{\mathbb{R}^n} A(|f(x)|) M_\varphi w dx. \quad (7.2)$$

Let us denote  $\mathcal{S}_k = \{Q \in \mathcal{S} : 4^{-k-1} < \|f\|_{A,Q} \leq 4^{-k}\}$  and set

$$T_k f(x) = \sum_{Q \in \mathcal{S}_k} \|f\|_{A,Q} \chi_Q(x).$$

If  $E \cap Q \neq \emptyset$  for some  $Q \in \mathcal{S}$  then we have that  $\|f\|_{A,Q} \leq \frac{1}{4}$  so necessarily

$$\mathcal{A}_{\mathcal{S},A} f(x) = \sum_{k=1}^{\infty} T_k f(x) \quad x \in E.$$

Since  $A$  is submultiplicative it satisfies 7.1 with  $\Lambda_A = A(4)$ . Using Lemma 8 with  $\mathcal{F}_k = \mathcal{S}_k$  combined with the fact that  $T_k f(x) \leq 4^{-k} \sum_{Q \in \mathcal{S}_k} \chi_Q(x)$  we have that

$$\int_E T_k f w dx \leq 2^{-k} w(E) + c \frac{4^{-k+1} A(4^k)}{\bar{\varphi}^{-1}((2\Lambda_A)^{2^k})} \int_{\mathbb{R}^n} A(|f|) M_\varphi w dx. \quad (7.3)$$

Taking that estimate into account,

$$\begin{aligned} w(E) &\leq \frac{1}{4} \int_E \mathcal{A}_{\mathcal{S}, A} f w dx \leq \frac{1}{4} \sum_{k=1}^{\infty} \int_E T_k f w dx \\ &\leq \frac{1}{4} w(E) + c \sum_{k=1}^{\infty} \frac{4^{-k} A(4^k)}{\bar{\varphi}^{-1}(2^{2^k})} \int_{\mathbb{R}^n} A(|f|) M_{\varphi} w dx. \end{aligned}$$

Now we observe that

$$\int_{2^{2^k-1}}^{2^{2^k}} \frac{1}{t \log(e+t)} dt \geq c. \quad (7.4)$$

Taking this into account, since  $\frac{A(t)}{t}$  is non-decreasing,

$$\begin{aligned} \sum_{k=1}^{\infty} \frac{4^{-k} A(4^k)}{\bar{\varphi}^{-1}(2^{2^k})} &\leq c \sum_{k=1}^{\infty} \int_{2^{2^k-1}}^{2^{2^k}} \frac{1}{t \log(e+t)} dt \frac{4^{-k} A(4^k)}{\bar{\varphi}^{-1}(2^{2^k})} \\ &\leq c \frac{A(4)}{4} \sum_{k=1}^{\infty} \int_{2^{2^k-1}}^{2^{2^k}} \frac{1}{t \bar{\varphi}^{-1}(t) \log(e+t)} dt \frac{A(4^{k-1})}{4^{k-1}} \\ &\leq c \frac{A(4)}{4} \sum_{k=1}^{\infty} \int_{2^{2^k-1}}^{2^{2^k}} \frac{A(\log(e+t)^2)}{t \bar{\varphi}^{-1}(t) \log(e+t) \log(e+t)^2} dt \\ &\leq c \int_1^{\infty} \frac{\varphi^{-1}(t) A(\log(e+t)^2)}{t^2 \log(e+t)^3} dt. \end{aligned}$$

This yields that that (7.2) holds with  $\kappa_{\varphi} = \int_1^{\infty} \frac{\varphi^{-1}(t) A(\log(e+t)^2)}{t^2 \log(e+t)^3} dt$ .

**7.2. Proof of theorem 8.** Taking into account Theorem 1 it suffices to obtain an endpoint estimate for each

$$\mathcal{A}_{\mathcal{S}}^{m,h}(b, f)(x) = \sum_{Q \in \mathcal{S}} |b(x) - b_Q|^{m-h} \|f|b - b_Q|^h\|_{B,Q} \chi_Q(x)$$

We shall consider two cases.

Assume first that  $h = m$ . Then we have that

$$\mathcal{A}_{\mathcal{S}}^{m,m}(b, f)(x) = \sum_{Q \in \mathcal{S}} \|f|b - b_Q|^m\|_{B,Q} \chi_Q(x) \leq \|b\|_{BMO}^m \sum_{Q \in \mathcal{S}} \|f\|_{A_m, Q} \chi_Q(x)$$

and it suffices to use the estimate obtained in Subsection 7.1, namely we have that

$$w \left( \left\{ x \in \mathbb{R}^n : \sum_{Q \in \mathcal{S}} \|f\|_{A_m, Q} \chi_Q(x) > \lambda \right\} \right) \leq c \kappa_{\varphi_m} \int_{\mathbb{R}^n} A_m \left( \frac{|f(x)|}{\lambda} \right) M_{\varphi_m} w(x) dx$$

where

$$\kappa_{\varphi_m} = \int_1^{\infty} \frac{\varphi_m^{-1}(t) A_m(\log(e+t)^2)}{t^2 \log(e+t)^3} dt.$$

Now we consider the case  $0 \leq h < m$ . Using generalized Hölder inequality if  $h > 0$  we have that

$$\mathcal{A}_{\mathcal{S}}^{m,h}(b, f)(x) \leq c \|b\|_{BMO}^h \sum_{Q \in \mathcal{S}} |b(x) - b_Q|^{m-h} \|f\|_{A_h, Q} \chi_Q(x) = \mathcal{T}_b^h f(x)$$

We define

$$E = \{x : |\mathcal{T}_b^h f(x)| > 8, M_{A_h} f(x) \leq 1/4\}.$$



By the Fefferman-Stein inequality (Lemma 3) and by homogeneity, it suffices to assume that  $\|b\|_{BMO} = 1$  and to show that

$$w(E) \leq cC_\varphi \int_{\mathbb{R}^n} A_h(|f|) M_{(\Phi_{m-h} \circ \varphi_h)(L)} w dx.$$

Let

$$\mathcal{S}_k = \{Q \in \mathcal{S} : 4^{-k-1} < \|f\|_{A_h, Q} \leq 4^{-k}\}$$

and for  $Q \in \mathcal{S}_k$ , set

$$F_k(Q) = \left\{ x \in Q : |b(x) - b_Q|^{m-h} > \left(\frac{3}{2}\right)^k \right\}.$$

If  $E \cap Q \neq \emptyset$  for some  $Q \in \mathcal{S}$ , then  $\|f\|_{A_h, Q} \leq 1/4$ . Therefore, for  $x \in E$ ,

$$\begin{aligned} |\mathcal{T}_b^h f(x)| &\leq \sum_{k=1}^{\infty} \sum_{Q \in \mathcal{S}_k} |b(x) - b_Q|^{m-h} \|f\|_{A_h, Q} \chi_Q(x) \\ &\leq \sum_{k=1}^{\infty} (3/2)^k \sum_{Q \in \mathcal{S}_k} \|f\|_{A_h, Q} \chi_Q(x) + \sum_{k=1}^{\infty} \sum_{Q \in \mathcal{S}_k} |b(x) - b_Q|^{m-h} \|f\|_{A_h, Q} \chi_{F_k(Q)}(x) \\ &\equiv \mathcal{T}_1 f(x) + \mathcal{T}_2 f(x). \end{aligned}$$

Let  $E_i = \{x \in E : \mathcal{T}_i f(x) > 4\}, i = 1, 2$ . Then

$$w(E) \leq w(E_1) + w(E_2). \quad (7.5)$$

Using (7.3) (with any Young function  $\psi_h$ )

$$\int_{E_1} (\mathcal{T}_1 f) w dx \leq \left( \sum_{k=1}^{\infty} (3/4)^k \right) w(E_1) + c_A \Lambda_A \sum_{k=1}^{\infty} \frac{(3/8)^k A_h(4^k)}{\overline{\psi_h}^{-1}(2^{2^k})} \int_{\mathbb{R}^n} A_h(|f|) M_{\psi_h} w dx.$$

This estimate, combined with  $w(E_1) \leq \frac{1}{4} \int_{E_1} (\mathcal{T}_1 f) w dx$ , implies

$$w(E_1) \leq c_A \Lambda_A \sum_{k=1}^{\infty} \frac{(3/8)^k A_h(4^k)}{\overline{\psi_h}^{-1}(2^{2^k})} \int_{\mathbb{R}^n} A_h(|f|) M_{\psi_h} w dx.$$

Now we observe that using (7.4)

$$\begin{aligned} \sum_{k=1}^{\infty} \frac{(3/8)^k A_h(4^k)}{\overline{\psi_h}^{-1}(2^{2^k})} &= \sum_{k=1}^{\infty} 2^k \frac{A_h(4^k)}{\overline{\psi_h}^{-1}(2^{2^k}) 4^k} \\ &\leq c \sum_{k=1}^{\infty} 2^k \frac{A_h(4^k)}{\overline{\psi_h}^{-1}(2^{2^k}) 4^k} \int_{2^{2^{k-1}}}^{2^{2^k}} \frac{1}{t \log(e+t)} dt \\ &\leq c \int_1^{\infty} \frac{\psi_h^{-1}(t) A_h(\log(e+t)^2)}{t^2 \log(e+t)^3} dt. \end{aligned}$$

We observe that since  $\frac{A_h(t)}{t}$  is not decreasing,

$$\frac{A_h(\log(e+t)^2)}{\log(e+t)^2} \leq \frac{A_h(\log(e+t)^{3(m-h)})}{\log(e+t)^{3(m-h)}} \leq \frac{A_h(\log(e+t)^{4(m-h)})}{\log(e+t)^{3(m-h)}},$$

we have that  $c \int_1^{\infty} \frac{\psi_h^{-1}(t) A_h(\log(e+t)^{4(m-h)})}{t^2 \log(e+t)^{3(m-h)}} dt$ , and choosing  $\psi_h = \Phi_{m-h} \circ \varphi_h$ ,

$$w(E_1) \leq c\kappa_h \int_{\mathbb{R}^n} A_h(|f|) M_{\Phi_{m-h} \circ \varphi_h} w dx$$

Now we focus on the estimate of  $w(E_2)$ . Arguing as in the proof of [24, Lemma 4.3], for  $Q \in \mathcal{S}_k$  we can define pairwise disjoint subsets  $E_Q \subseteq Q$  and prove that

$$1 \leq \frac{c}{|Q|} \int_{E_Q} A_h(4^k |f|) dx.$$

Hence,

$$w(E_2) \leq \frac{1}{4} \|\mathcal{T}_2 f\|_{L^1} c \sum_{k=1}^{\infty} \sum_{Q \in \mathcal{S}_k} \frac{1}{4^k} \left( \frac{1}{|Q|} \int_{F_k(Q)} |b - b_Q|^{m-h} w dx \right) \int_{E_Q} A_h(4^k |f|) dx. \quad (7.6)$$

Now we apply twice the generalized Hölder inequality (3.2). First we obtain the following inequality

$$\frac{1}{|Q|} \int_{F_k(Q)} |b - b_Q|^{m-h} w dx \leq c_n \|w \chi_{F_k(Q)}\|_{L(\log L)^{m-h}, Q}. \quad (7.7)$$

Now we define  $\Phi_{m-h}(t) = t \log(e + t)^{m-h}$ , and  $\Psi_{m-h}$  as

$$\Psi_{m-h}^{-1}(t) = \frac{\Phi_{m-h}^{-1}(t)}{\varphi_h^{-1} \circ \Phi_{m-h}^{-1}(t)}.$$

Since  $\varphi_h(t)/t$  and  $\Phi$  are strictly increasing functions,  $\Psi_{m-h}$  is strictly increasing, too. Hence, a direct application of (3.5) yields

$$\begin{aligned} \|w \chi_{F_k(Q)}\|_{L(\log L)^{m-h}, Q} &\leq 2 \|\chi_{F_k(Q)}\|_{\Psi, Q} \|w\|_{(\Phi_{m-h} \circ \varphi_h), Q} \\ &= \frac{2}{\Psi_{m-h}^{-1}(|Q|/|F_k(Q)|)} \|w\|_{(\Phi_{m-h} \circ \varphi_h), Q}. \end{aligned} \quad (7.8)$$

Now we observe that Theorem 14 assures that  $|F_k(Q)| \leq \alpha_k |Q|$ , where  $\alpha_k = \min(1, e^{-\frac{(3/2)\frac{k}{m-h}}{2^k e} + 1})$ . That fact together with (7.7) and (7.8) yields

$$\frac{1}{|Q|} \int_{F_k(Q)} |b - b_Q|^j w dx \leq \frac{c_n}{\Psi_{m-h}^{-1}(1/\alpha_k)} \|w\|_{(\Phi_{m-h} \circ \varphi_h), Q}.$$

From this estimate combined with (7.6) it follows that

$$\begin{aligned} w(E_2) &\leq c_n \sum_{k=1}^{\infty} \frac{1}{\Psi_{m-h}^{-1}(1/\alpha_k) 4^k} \sum_{Q \in \mathcal{S}_k} \|w\|_{(\Phi_{m-h} \circ \varphi_h), Q} \int_{E_Q} A_h(4^k |f|) dx \\ &\leq c_n \left( \sum_{k=1}^{\infty} \frac{1}{\Psi_{m-h}^{-1}(1/\alpha_k)} \frac{A_h(4^k)}{4^k} \right) \int_{\mathbb{R}^n} A_h(|f|) M_{(\Phi_{m-h} \circ \varphi_h)(L)} w(x) dx. \end{aligned}$$

Now we observe that we can choose  $c_{n,m,h}$  such that for every  $k > c_{n,m,h}$  we have that  $\frac{1}{\alpha_{k-1}} = e^{\frac{(3/2)\frac{k-1}{m-h}}{2^{k-1} e} - 1} \geq \max\{e^2, 4^k\}$ . We note that

$$\int_{\frac{1}{\alpha_{k-1}}}^{\frac{1}{\alpha_k}} \frac{1}{t \log(e + t)} dt \geq c.$$

Taking this into account, if  $\frac{1}{\beta} = (m-h)\frac{\log 4}{\log(3/2)}$ , since  $A$  is submultiplicative and  $\frac{A(t)}{t}$  is non-decreasing, we obtain

$$\begin{aligned} \sum_{k=1}^{\infty} \frac{1}{\Psi_{m-h}^{-1}(1/\alpha_k)} \frac{A_h(4^k)}{4^k} &\leq \alpha_{n,h,m} + \sum_{k=c_{n,m,h}}^{\infty} \frac{1}{\Psi_{m-h}^{-1}(1/\alpha_k)} \frac{A_h(4^k)}{4^k} \\ &\leq \alpha_{n,h,m} + c_n \frac{A(4)}{4} \int_1^{\infty} \frac{1}{\Psi_{m-h}^{-1}(t)} \frac{1}{t \log(e+t)} \frac{A_h(\log(e+t)^{1/\beta})}{\log(e+t)^{1/\beta}} dt \\ &\leq \alpha_{n,h,m} + c_n \int_1^{\infty} \frac{\varphi_h^{-1} \circ \Phi_{m-h}^{-1}(t)}{\Phi_{m-h}^{-1}(t)} \frac{1}{t \log(e+t)} \frac{A_h(\log(e+t)^{4(m-h)})}{\log(e+t)^{4(m-h)}} dt \\ &\simeq \alpha_{n,h,m} + c_n \int_1^{\infty} \frac{\varphi_h^{-1} \circ \Phi_{m-h}^{-1}(t) A_h(\log(e+t)^{4(m-h)})}{t^2 \log(e+t)^{3(m-h)+1}} dt \end{aligned}$$

## 8. PROOFS OF EXPONENTIAL DECAY ESTIMATES

**8.1. Proof of Theorem 9.** We recall that in [34, Theorem 2.1], it was established that

$$\left| \left\{ x \in Q : \sum_{R \in \mathcal{S}, R \subseteq Q} \chi_R(x) > t \right\} \right| \leq ce^{-\alpha t} |Q| \quad (8.1)$$

Assume that  $\text{supp } f \subset Q_0$ . It is easy to see that (4.5) holds with  $b_{R_Q}$  replaced by  $b_{3Q}$ . Then we have that for almost every  $x \in Q_0$

$$|T_b^m(f)(x)| = |T_b^m(f\chi_{3Q})(x)| \leq c_{n,m} c_T \sum_{h=0}^m \mathcal{C}_{B,\mathcal{F}}^{m,h}(b, f)$$

where

$$\mathcal{C}_{B,\mathcal{F}}^{m,h}(b, f) = \sum_{Q \in \mathcal{F}} |b(x) - b_{3Q}|^{m-h} \|f\|_{B,3Q}^h \chi_Q(x)$$

and  $\mathcal{S} \subset \mathcal{D}(Q_0)$  is a sparse family. For the sake of clarity we consider now two cases. If  $m = 0$  then we only have to deal with  $\mathcal{C}_{B,\mathcal{F}}^{0,0}(b, f) = \sum_{Q \in \mathcal{F}} \|f\|_{B,3Q} \chi_Q(x)$ . In this case taking into account that

$$\frac{\sum_{Q \in \mathcal{F}} \|f\|_{B,3Q} \chi_Q(x)}{M_B f(x)} \leq \sum_{Q \in \mathcal{F}} \chi_Q(x)$$

a direct application of (8.1) yields (1.12).

For the case  $m > 0$ . First we observe that

$$|b(x) - b_{3Q}|^{m-h} \leq c_{n,m} \|b\|_{BMO}^{m-h} + c_{n,m} |b(x) - b_Q|^{m-h}$$

and also that by generalized Hölder inequality and taking into account (3.5) and (3.6),

$$\| |b - b_{3Q}|^h f \|_{B,3Q} \leq \|b\|_{BMO}^h \|f\|_{A,3Q}.$$

Then we have that

$$\begin{aligned} &\left| \left\{ x \in Q_0 : \frac{\mathcal{A}_{B,\mathcal{F}}^{m,h}(b, f)}{M_A f} > \lambda \right\} \right| \\ &\leq \left| \left\{ x \in Q_0 : \frac{\sum_{Q \in \mathcal{F}} \|f\|_{A,3Q} \chi_Q(x)}{M_A f} > \frac{\lambda}{2c_{n,m} \|b\|_{BMO}^m c_T} \right\} \right| \\ &+ \left| \left\{ x \in Q_0 : \frac{\sum_{Q \in \mathcal{F}} |b(x) - b_Q|^{m-h} \|f\|_{A,3Q} \chi_Q(x)}{M_A f} > \frac{\lambda}{2c_{n,m} \|b\|_{BMO}^h c_T} \right\} \right| \\ &= I + II. \end{aligned}$$

For  $I$  we observe that

$$\frac{\sum_{Q \in \mathcal{F}} \|f\|_{A,3Q} \chi_Q(x)}{M_A f} \leq \sum_{Q \in \mathcal{S}} \chi_Q(x)$$

and then a direct application of (8.1) yields

$$\left| \left\{ x \in Q_0 : \frac{\sum_{Q \in \mathcal{S}} \|f\|_{A,3Q} \chi_Q(x)}{M_A f} > \frac{\lambda}{2c_{n,m} \|b\|_{BMO}^m} \right\} \right| \leq c e^{-\alpha \frac{\lambda}{2c_{n,m} \|b\|_{BMO}^m}} |Q|$$

Now we focus on  $II$ . [24, Lemma 5.1] provides a sparse family  $\tilde{\mathcal{F}}$  such that for every  $Q \in \mathcal{F}$ ,

$$|b(x) - b_Q| \leq c_n \sum_{P \in \tilde{\mathcal{F}}, P \subseteq Q} \left( \frac{1}{|P|} \int_P |b(x) - b_P| dx \right) \chi_P(x).$$

Since  $b \in BMO$  we have that for every  $Q \in \mathcal{F}$ ,

$$|b(x) - b_Q| \leq c_n \sum_{P \in \tilde{\mathcal{F}}, P \subseteq Q} \left( \frac{1}{|P|} \int_P |b(x) - b_P| dx \right) \chi_P(x) \leq c_n \|b\|_{BMO} \sum_{P \in \tilde{\mathcal{F}}, P \subseteq Q_0} \chi_P(x).$$

Then we have that

$$\begin{aligned} \frac{\sum_{Q \in \mathcal{S}} |b(x) - b_Q|^{m-h} \|f\|_{A,3Q} \chi_Q(x)}{M_A f} &\leq c_{n,m,h} \|b\|_{BMO}^{m-h} \sum_{Q \in \mathcal{F}} \left( \sum_{P \in \tilde{\mathcal{F}}, P \subseteq Q_0} \chi_P(x) \right)^{m-h} \chi_Q(x) \\ &\leq c_{n,m,h} \|b\|_{BMO}^{m-h} \left( \sum_{P \in \tilde{\mathcal{F}}, P \subseteq Q_0} \chi_P(x) \right)^{m-h+1} \chi_Q(x) \end{aligned}$$

and taking using again (8.1),

$$\begin{aligned} II &\leq \left| \left\{ x \in Q_0 : c_{n,m,h} \|b\|_{BMO}^{m-h} \left( \sum_{P \in \tilde{\mathcal{F}}, P \subseteq Q_0} \chi_P(x) \right)^{m-h+1} > \frac{\lambda}{2c_{n,m} \|b\|_{BMO}^h} \right\} \right| \\ &= \left| \left\{ x \in Q_0 : c_{n,m,h} \sum_{P \in \tilde{\mathcal{F}}, P \subseteq Q_0} \chi_P(x) > \left( \frac{\lambda}{2c_{n,m} \|b\|_{BMO}^m} \right)^{\frac{1}{m-h+1}} \right\} \right| \leq c e^{-\alpha \left( \frac{\lambda}{2c_{n,m} \|b\|_{BMO}^m} \right)^{\frac{1}{m-h+1}}} |Q| \end{aligned}$$

as we wanted to prove. Controlling all the decays by the worst possible, namely, when  $h = 0$  we are done.

## 9. PROOFS OF APPLICATIONS

**9.1. Proof of Theorem 10.** Since  $T$  is an  $\omega$ -Calderón-Zygmund operator, we know that it satisfies an  $L^\infty$ -Hörmander with  $H_\infty \leq c_n (\|\omega\|_{\text{Dini}} + c_K)$  condition, then  $A_0(t) = t$ . Let us call  $\Phi_j(t) = t \log(e+t)^j$ . We are going to apply Theorem 8 with  $A_j(t) = \Phi_j(t)$ , so we have to make suitable choices for each  $\varphi_h$  to obtain the desired estimate for each term

$$\kappa_{\varphi_h} \int_{\mathbb{R}^n} A_h \left( \frac{|f(x)|}{\lambda} \right) M_{\Phi_{m-h} \circ \varphi_h} w(x) dx.$$

We consider three cases. Let us assume first that  $0 < h < m$ . Then

$$\begin{aligned}
\kappa_{\varphi_h} &= \alpha_{n,m,h} + c_n \int_1^\infty \frac{\varphi_h^{-1} \circ \Phi_{m-h}^{-1}(t) A_h(\log(e+t)^{4(m-h)})}{t^2 \log(e+t)^{3(m-h)+1}} dt \\
&\lesssim \alpha_{n,m,h} + c_n \int_1^\infty \frac{\varphi_h^{-1}(t) \log(e + \log(e + \Phi_{m-h}(t))^{4(m-h)})^h}{\Phi_{m-h}(t)^2 \log(e + \Phi_{m-h}(t))^{1-(m-h)}} \Phi'_{m-h}(t) dt \\
&\lesssim \alpha_{n,m,h} + c_n \int_1^\infty \frac{\varphi_h^{-1}(t) \log(e + \log(e + \Phi_{m-h}(t))^{4(m-h)})^h}{t \Phi_{m-h}(t) \log(e + \Phi_{m-h}(t))^{1-(m-h)}} dt \\
&\lesssim \alpha_{n,m,h} + c_n \int_1^\infty \frac{\varphi_h^{-1}(t) \log(e + \log(e + \Phi_{m-h}(t))^{4(m-h)})^h}{t^2 \log(e+t)} dt.
\end{aligned}$$

If we choose  $\varphi_h(t) = t \log(e+t) \log(e + \log(e+t))^{1+\epsilon}$ ,  $\epsilon > 0$ , then

$$\begin{aligned}
\kappa_{\varphi_h} &\lesssim \alpha_{n,m,h} + c_n \int_1^\infty \frac{\log(e + \log(e + \Phi_{m-h}(t))^{4(m-h)})^h}{t \log(e+t)^2 \log(e + \log(e+t))^{1+\epsilon}} dt \\
&\lesssim \alpha_{n,m,h} + c_n \int_1^\infty \frac{dt}{t \log(e+t) \log(e + \log(e+t))^{1+\epsilon}} \\
&\lesssim \frac{1}{\epsilon}
\end{aligned}$$

and we observe that also

$$\Phi_{m-h} \circ \varphi_h \lesssim t \log(e+t)^m \log(e + \log(e+t))^{1+\epsilon}. \quad (9.1)$$

Then for  $0 < h < m$

$$\kappa_{\varphi_h} \int_{\mathbb{R}^n} A_h \left( \frac{|f(x)|}{\lambda} \right) M_{\Phi_{m-h} \circ \varphi_h} w(x) dx \leq c \frac{1}{\epsilon} \int_{\mathbb{R}^n} \Phi_m \left( \frac{|f(x)|}{\lambda} \right) M_{L(\log L)^m (\log \log L)^{1+\epsilon}} w(x) dx$$

For the case  $h = 0$ , arguing as in the first case, we obtain

$$\begin{aligned}
\kappa_{\varphi_0} &= \alpha_{n,m} + c_n \int_1^\infty \frac{\varphi_0^{-1} \circ \Phi_m^{-1}(t) A_0(\log(e+t)^{4m})}{t^2 \log(e+t)^{3m+1}} dt \\
&\lesssim \alpha_{n,m} + c_n \int_1^\infty \frac{\varphi_0^{-1}(t)}{t^2 \log(e+t)} dt
\end{aligned}$$

So it suffices to choose  $\varphi_0(t) = t \log(e + \log(e+t))^{1+\epsilon}$  and have that  $\kappa_{\varphi_0} < \frac{1}{\epsilon}$  and

$$\Phi_m \circ \varphi_0 \lesssim \varphi_0(t) \log(e+t)^m = t \log(e+t)^m \log(e + \log(e+t))^{1+\epsilon}. \quad (9.2)$$

Consequently

$$\kappa_{\varphi_0} \int_{\mathbb{R}^n} A_0 \left( \frac{|f(x)|}{\lambda} \right) M_{\Phi_m \circ \varphi_0} w(x) dx \leq c \frac{1}{\epsilon} \int_{\mathbb{R}^n} \frac{|f(x)|}{\lambda} M_{L(\log L)^m (\log \log L)^{1+\epsilon}} w(x) dx.$$

To end the proof we consider  $h = m$ . We observe that

$$\begin{aligned}
\kappa_{\varphi_m} &= \int_1^\infty \frac{\varphi_m^{-1}(t) A_m(\log(e+t)^2)}{t^2 \log(e+t)^3} dt \\
&= \int_1^\infty \frac{\varphi_m^{-1}(t) \log(e + \log(e+t)^2)^m}{t^2 \log(e+t)} dt
\end{aligned}$$

and taking  $\varphi_m(t) = t \log(e+t)^m \log(e + \log(e+t))^{1+\epsilon}$ , we obtain  $\kappa_{\varphi_m} < \frac{1}{\epsilon}$  and since  $\Phi_0(t) = t$

$$\kappa_{\varphi_m} \int_{\mathbb{R}^n} A_m \left( \frac{|f(x)|}{\lambda} \right) M_{\Phi_0 \circ \varphi_m} w(x) dx \leq c \frac{1}{\epsilon} \int_{\mathbb{R}^n} \Phi_m \left( \frac{|f(x)|}{\lambda} \right) M_{L(\log L)^m (\log \log L)^{1+\epsilon}} w(x) dx$$

Collecting the preceding estimates

$$\begin{aligned} w(\{x \in \mathbb{R}^n : T_b^m f(x) > \lambda\}) &\leq c_n C_T \sum_{h=0}^m \left( \kappa_{\varphi_h} \int_{\mathbb{R}^n} A_h \left( \frac{|f(x)|}{\lambda} \right) M_{\Phi_{m-h} \circ \varphi_h} w(x) dx \right) \\ &\leq c_{n,m} C_T \frac{1}{\varepsilon} \int_{\mathbb{R}^n} \Phi_m \left( \frac{|f(x)|}{\lambda} \right) M_{L(\log L)^m (\log \log L)^{1+\varepsilon}} w(x) dx. \end{aligned}$$

Now we observe that since  $t \log(e+t)^m \log(e + \log(e+t))^{1+\varepsilon} \leq ct \log(e+t)^{m+\varepsilon}$  for  $t \geq 1$  we also have that

$$w(\{x \in \mathbb{R}^n : T_b^m f(x) > \lambda\}) \leq c_{n,m} C_T \frac{1}{\varepsilon} \int_{\mathbb{R}^n} \Phi_m \left( \frac{|f(x)|}{\lambda} \right) M_{L(\log L)^{m+\varepsilon}} w(x) dx.$$

Now we turn our attention now to the remaining estimates. Assume that  $w \in A_\infty$ . To prove (2.2) we argue as in [17, Corollary 1.4]. Since  $\log(t) \leq \frac{t^\alpha}{\alpha}$ , for every  $t \geq 1$  we have that

$$\frac{1}{\varepsilon} M_{L(\log L)^{m+\varepsilon}} w \leq c \frac{1}{\varepsilon} \frac{1}{\alpha^{m+\varepsilon}} M_{1+(m+\varepsilon)\alpha} w.$$

Taking  $(m+\varepsilon)\alpha = \frac{1}{\tau_n [w]_{A_\infty}}$  where  $\tau_n$  is chosen as in Lemma 4 we have that, precisely, using Lemma 4,

$$\frac{1}{\varepsilon} \frac{1}{\alpha^\varepsilon} M_{1+(m+\varepsilon)\alpha} w = \frac{1}{\varepsilon} ((m+\varepsilon)\tau_n \varepsilon [w]_{A_\infty})^{m+\varepsilon} M_{1+\frac{1}{\tau_n [w]_{A_\infty}}} w \leq c_m \frac{1}{\varepsilon} [w]_{A_\infty}^{m+\varepsilon} M w.$$

Finally choosing  $\varepsilon = \frac{1}{\log(e+[w]_{A_\infty})}$  we have that

$$\frac{1}{\varepsilon} M_{L(\log L)^{m+\varepsilon}} w \leq c_m \frac{1}{\varepsilon} [w]_{A_\infty}^{m+\varepsilon} M w \leq c_m \log(e+[w]_{A_\infty}) [w]_{A_\infty}^m M w.$$

This estimate combined with (2.1) yields (2.2). We end the proof noting that (2.3) follows from (2.2) and the definition of  $w \in A_1$ .

**9.2. Proof of Theorem 11.** It suffices to prove that  $K \in \mathcal{H}_{\overline{B}}$ , namely that  $T$  is a  $\overline{B}$ -Hörmander operator. The rest of the statements of the Theorem follow from applying the main Theorems to  $T_\Omega$ . Let us prove then that  $K \in \mathcal{H}_{\overline{B}}$ . We borrow the following estimate from [29, Proposition 4.2],

$$\|K(\cdot - y) - K(\cdot)\|_{\overline{B}, s \leq |x| < 2s} \leq cs^{-n} \left( \frac{|y|}{s} + \omega_{\overline{B}} \left( \frac{|y|}{s} \right) \right), \quad |y| < \frac{s}{2}.$$

This condition is essentially equivalent to consider cubes instead of balls, and hence to our condition. We also note that in the convolution case it suffices to consider balls centered at the origin.

Now we observe that choosing  $s = 2^k R$  and taking  $|y| < R \leq \frac{s}{2}$  we have that

$$\begin{aligned} \sum_{k=1}^{\infty} (2^k R)^n \|K(\cdot - y) - K(\cdot)\|_{\overline{B}, 2^k R \leq |x| < 2^{k+1} R} &\leq c \left( \sum_{k=1}^{\infty} 2^{-k} + \omega_{\overline{B}}(2^{-k}) \right) \\ &\leq c + c \int_0^1 w_{\overline{B}}(t) \frac{1}{t} dt. \end{aligned}$$

Hence taking into account (2.4) we have that  $K \in \mathcal{H}_{\overline{B}}$ .

**9.3. Proof of Theorem 12.** First we check first that both (1.7) and (1.8) hold. Let us choose  $r' = \frac{n}{l} + \varepsilon$ . Lemma 1 yields then that  $K^N \in \mathcal{H}_{L^r(\log L)^{mr}}$ . Let us call  $T_N$  the truncation of  $T$  associated to  $K_N$ . For the case  $m = 0$  we deal with  $T$  and we have that  $K_N \in \mathcal{H}_{L^r}$  so it suffices to apply Theorem 4 with  $\overline{B}(t) = t^r$  to each  $T_N$  and apply a standard approximation argument. For the case  $m > 0$ , let us call  $\overline{B}_m(t) = t^r \log(e+t)^{mr}$ . We choose  $A(t) = t^{r'}$  so we have that  $A^{-1}(t) \overline{B}^{-1}(t) \overline{C}_m^{-1}(t) \leq ct$  for every  $t \geq 1$  where  $\overline{C}_m(t) = e^{t^{1/m}}$ . Then (1.8) holds for  $T_N$  and any  $b \in BMO$  with constant independent of  $N$  and a standard approximation argument yields that those estimates hold.

Now we turn our attention to the strong type estimate. We observe that it also follows from 1 that  $K_N$  satisfies an  $A$ -Hörmander condition with  $A(t) = t^r$  and that  $\mathcal{K}_{r,A} = 1$ . Then we can

apply Theorem 2 to each  $T_N$  and the desired estimate follows again from a standard approximation argument.

#### APPENDIX A. QUANTITATIVE UNWEIGHTED ESTIMATES

In this appendix we gather some precise unweighted estimates that we will need to obtain the fully quantitative sparse domination in Theorem 1.

**Lemma 9.** *Let  $S$  be a linear operator such that  $S : L^1(\mu) \rightarrow L^{1,\infty}(\mu)$  and  $\nu \in (0, 1)$ . Then if  $E$  is a measurable set such that  $0 < \mu(E) < \infty$*

$$\int_E |Sf(x)|^\nu d\mu \leq 2 \frac{\nu}{1-\nu} \|S\|_{L^1 \rightarrow L^{1,\infty}}^\nu \mu(E)^{1-\nu} \|f\|_{L^1}^\nu.$$

*Proof.* It suffices to track constants in [11, Lemma 5.6] choosing  $C = \|S\|_{L^1 \rightarrow L^{1,\infty}}$ .  $\square$

We are not aware of the appearance of the following result in the literature. It essentially allows us to interpolate between  $L^p$  scales to obtain a modular inequality and it will be fundamental to obtain a suitable control for  $\mathcal{M}_T$  in Lemma 7.

**Lemma 10.** *Let  $A$  be a Young function such that  $A \in \mathcal{Y}(p_0, p_1)$ . Let  $G$  a sublinear operator of weak type  $(p_0, p_0)$  and of weak type  $(p_1, p_1)$ . Then*

$$|\{x \in \mathbb{R}^n : |G(x)| > t\}| \leq \int_{\mathbb{R}^n} A\left(\kappa \frac{|f(x)|}{t}\right) dx$$

where  $\kappa = 2 \max\{\|G\|_{L^{p_0} \rightarrow L^{p_0,\infty}}, \|G\|_{L^{p_1} \rightarrow L^{p_1,\infty}}\}$

*Proof.* Let us call  $\kappa = 2 \max\{\|G\|_{L^{p_0} \rightarrow L^{p_0,\infty}}, \|G\|_{L^{p_1} \rightarrow L^{p_1,\infty}}\}$  and consider  $f(x) = f_1(x) + f_2(x)$  where

$$\begin{aligned} f_0(x) &= f(x) \chi_{\{|f(x)| > \frac{1}{\kappa} t_0 \lambda\}}(x), \\ f_1(x) &= f(x) \chi_{\{|f(x)| \leq \frac{1}{\kappa} t_0 \lambda\}}(x). \end{aligned}$$

Using the partition of  $f$  and the assumptions on  $G$  we have that

$$\begin{aligned} &|\{x \in \mathbb{R}^n : |Gf(x)| > \lambda\}| \\ &\leq \left| \left\{ x \in \mathbb{R}^n : |Gf_0(x)| > \frac{\lambda}{2} \right\} \right| + \left| \left\{ x \in \mathbb{R}^n : |Gf_1(x)| > \frac{\lambda}{2} \right\} \right| \\ &\leq 2^{p_0} \|G\|_{L^{p_0} \rightarrow L^{p_0,\infty}}^{p_0} \int_{\mathbb{R}^n} \left( \frac{|f_0(x)|}{\lambda} \right)^{p_0} dx + 2^{p_1} \|G\|_{L^{p_1} \rightarrow L^{p_1,\infty}}^{p_1} \int_{\mathbb{R}^n} \left( \frac{|f_1(x)|}{\lambda} \right)^{p_1} dx \\ &\leq \int_{\mathbb{R}^n} \left( \kappa \frac{|f_0(x)|}{\lambda} \right)^{p_0} dx + \int_{\mathbb{R}^n} \left( \kappa \frac{|f_1(x)|}{\lambda} \right)^{p_1} dx \end{aligned}$$

Now we observe that, using the hypothesis on  $A$ ,

$$\int_{\mathbb{R}^n} \left( \kappa \frac{|f_0(x)|}{\lambda} \right)^{p_0} dx = \int_{\{|f(x)| > \frac{1}{\kappa} t_0 \lambda\}} \left( \kappa \frac{|f(x)|}{\lambda} \right)^{p_0} dx \leq c_0 \int_{\{|f(x)| > \frac{1}{\kappa} t_0 \lambda\}} A\left(\kappa \frac{|f(x)|}{\lambda}\right) dx$$

and analogously

$$\int_{\mathbb{R}^n} \left( \kappa \frac{|f_1(x)|}{\lambda} \right)^{p_1} dx = \int_{\{|f(x)| \leq \frac{1}{\kappa} t_0 \lambda\}} \left( \kappa \frac{|f(x)|}{\lambda} \right)^{p_1} dx \leq c_1 \int_{\{|f(x)| \leq \frac{1}{\kappa} t_0 \lambda\}} A\left(\kappa \frac{|f(x)|}{\lambda}\right) dx$$

Then we have that combining those estimates and taking into account that since  $A$  is a Young function  $cA(t) \leq A(ct)$  for  $c \geq 1$ ,

$$|\{x \in \mathbb{R}^n : |Gf(x)| > \lambda\}| \leq \int_{\mathbb{R}^n} A\left(\kappa \frac{|f(x)|}{\lambda}\right) dx$$

$\square$

**Lemma 11.** *Let  $A$  be a Young function. If  $T$  is a  $\overline{A}$ -Hörmander operator then*

$$\|T\|_{L^1 \rightarrow L^{1,\infty}} \leq c_n (\|T\|_{L^2 \rightarrow L^2} + H_{\overline{A}})$$

*and as a consequence of Marcinkiewicz theorem and the fact that  $T$  is almost self-dual*

$$\|T\|_{L^p \rightarrow L^p} \leq c_n (\|T\|_{L^2 \rightarrow L^2} + H_{\overline{A}}).$$

*Proof.* For the endpoint estimate, following ideas in [19, Theorem A.1] it suffices to follow the standard proof using Hörmander condition, see for instance [11, Theorem 5.10], but with the following small twist in the argument. When estimating the level set  $\{|Tf(x)| > \lambda\}$  the Calderón-Zygmund decomposition of  $f$  has to be taken at level  $\alpha\lambda$  and optimize  $\alpha$  at the end of the proof.

For the strong type estimate it suffices to use the endpoint estimate we have just obtained combined with the  $L^2$  boundedness of the operator to obtain the corresponding bound in the range  $1 < p \leq 2$  and duality for the rest of the range.  $\square$

#### ACKNOWLEDGMENTS

The first author would like Carlos Pérez for inviting him to visit BCAM between January and April 2016, and BCAM for the warm hospitality shown during his visit.

#### REFERENCES

- [1] The Anh Bui, José M. Conde-Alonso, Xuan Thanh Duong, and Mahdi Hormozi. A note on weighted bounds for singular operators with nonsmooth kernels. *Studia Math.*, 236(3):245–269, 2017.
- [2] Marcela Caldarelli, Andrei K. Lerner, and Sheldy Ombrosi. On a counterexample related to weighted weak type estimates for singular integrals. *ArXiv e-prints*, June 2015.
- [3] Daewon Chung, M. Cristina Pereyra, and Carlos Pérez. Sharp bounds for general commutators on weighted Lebesgue spaces. *Trans. Amer. Math. Soc.*, 364(3):1163–1177, 2012.
- [4] Ronald R. Coifman, Richard Rochberg, and Guido Weiss. Factorization theorems for Hardy spaces in several variables. *Ann. of Math. (2)*, 103(3):611–635, 1976.
- [5] José M. Conde-Alonso, Amalia Culiuc, Francesco Di Plinio, and Yumeng Ou. A sparse domination principle for rough singular integrals. *ArXiv e-prints*, December 2016.
- [6] José M. Conde-Alonso and Guillermo Rey. A pointwise estimate for positive dyadic shifts and some applications. *Math. Ann.*, 365(3-4):1111–1135, 2016.
- [7] David Cruz-Uribe, José M. Martell, and Carlos Pérez. Extrapolation from  $A_\infty$  weights and applications. *J. Funct. Anal.*, 213(2):412–439, 2004.
- [8] David Cruz-Uribe, José M. Martell, and Carlos Pérez. Sharp weighted estimates for classical operators. *Adv. Math.*, 229(1):408–441, 2012.
- [9] David V. Cruz-Uribe, José M. Martell, and Carlos Pérez. *Weights, extrapolation and the theory of Rubio de Francia*, volume 215 of *Operator Theory: Advances and Applications*. Birkhäuser/Springer Basel AG, Basel, 2011.
- [10] Carlos Domingo-Salazar, Michael T. Lacey, and Guillermo Rey. Borderline weak-type estimates for singular integrals and square functions. *Bull. Lond. Math. Soc.*, 48(1):63–73, 2016.
- [11] Javier Duoandikoetxea. *Fourier analysis*, volume 29 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2001. Translated and revised from the 1995 Spanish original by David Cruz-Uribe.
- [12] Nobuhiko Fujii. Weighted bounded mean oscillation and singular integrals. *Math. Japon.*, 22(5):529–534, 1977/78.
- [13] Tuomas P. Hytönen. The sharp weighted bound for general Calderón-Zygmund operators. *Ann. of Math. (2)*, 175(3):1473–1506, 2012.
- [14] Tuomas P. Hytönen. The Holmes-Wick theorem on two-weight bounds for higher order commutators revisited. *Arch. Math. (Basel)*, 107(4):389–395, 2016.
- [15] Tuomas P. Hytönen and Kangwei Li. Weak and strong  $A_p$ - $A_\infty$  estimates for square functions and related operators. *ArXiv e-prints*, September 2015.
- [16] Tuomas P. Hytönen and Carlos Pérez. Sharp weighted bounds involving  $A_\infty$ . *Anal. PDE*, 6(4):777–818, 2013.
- [17] Tuomas P. Hytönen and Carlos Pérez. The  $L(\log L)^\epsilon$  endpoint estimate for maximal singular integral operators. *J. Math. Anal. Appl.*, 428(1):605–626, 2015.
- [18] Tuomas P. Hytönen, Carlos Pérez, and Ezequiel Rela. Sharp reverse Hölder property for  $A_\infty$  weights on spaces of homogeneous type. *J. Funct. Anal.*, 263(12):3883–3899, 2012.
- [19] Tuomas P. Hytönen, Luz Roncal, and Olli Tapiola. Quantitative weighted estimates for rough homogeneous singular integrals. *Israel J. Math.*, 218(1):133–164, 2017.
- [20] Michael T. Lacey. An elementary proof of the  $A_2$  bound. *Israel J. Math.*, 217(1):181–195, 2017.
- [21] Andrei K. Lerner. A simple proof of the  $A_2$  conjecture. *Int. Math. Res. Not. IMRN*, (14):3159–3170, 2013.



- [22] Andrei K. Lerner. On pointwise estimates involving sparse operators. *New York J. Math.*, 22:341–349, 2016.
- [23] Andrei K. Lerner and Fedor Nazarov. Intuitive dyadic calculus: the basics. *ArXiv e-prints*, August 2015.
- [24] Andrei K. Lerner, Sheldy Ombrosi, and Israel P. Rivera-Ríos. On pointwise and weighted estimates for commutators of Calderón-Zygmund operators. *ArXiv e-prints*, April 2016.
- [25] Kangwei Li. Personal communication.
- [26] Kangwei Li. Sparse domination theorem for multilinear singular integral operators with  $L^r$ -Hörmander condition. *ArXiv e-prints*, June 2016.
- [27] Liguang Liu and Teresa Luque. A  $B_p$  condition for the strong maximal function. *Trans. Amer. Math. Soc.*, 366(11):5707–5726, 2014.
- [28] María Lorente, José M. Martell, Carlos Pérez, and María S. Riveros. Generalized Hörmander conditions and weighted endpoint estimates. *Studia Math.*, 195(2):157–192, 2009.
- [29] María Lorente, José M. Martell, María S. Riveros, and Alberto de la Torre. Generalized Hörmander’s conditions, commutators and weights. *J. Math. Anal. Appl.*, 342(2):1399–1425, 2008.
- [30] María Lorente, María S. Riveros, and Alberto de la Torre. Weighted estimates for singular integral operators satisfying Hörmander’s conditions of Young type. *J. Fourier Anal. Appl.*, 11(5):497–509, 2005.
- [31] José M. Martell, Carlos Pérez, and Rodrigo Trujillo-González. Lack of natural weighted estimates for some singular integral operators. *Trans. Amer. Math. Soc.*, 357(1):385–396, 2005.
- [32] Richard O’Neil. Fractional integration in Orlicz spaces. I. *Trans. Amer. Math. Soc.*, 115:300–328, 1965.
- [33] Carmen Ortiz-Caraballo. *Conmutadores de integrales singulares y pesos  $A_1$* . PhD thesis, Universidad de Sevilla, 2011.
- [34] Carmen Ortiz-Caraballo, Carlos Pérez, and Ezequiel Rela. Exponential decay estimates for singular integral operators. *Math. Ann.*, 357(4):1217–1243, 2013.
- [35] C. Pérez. On sufficient conditions for the boundedness of the Hardy-Littlewood maximal operator between weighted  $L^p$ -spaces with different weights. *Proc. London Math. Soc. (3)*, 71(1):135–157, 1995.
- [36] Carlos Pérez. Endpoint estimates for commutators of singular integral operators. *J. Funct. Anal.*, 128(1):163–185, 1995.
- [37] Carlos Pérez and Gladis Pradolini. Sharp weighted endpoint estimates for commutators of singular integrals. *Michigan Math. J.*, 49(1):23–37, 2001.
- [38] Carlos Pérez and Israel P. Rivera-Ríos. Borderline weighted estimates for commutators of singular integrals. *Israel J. Math.*, 217(1):435–475, 2017.
- [39] Carlos Pérez and Israel P. Rivera-Ríos. Three observations on commutators of Singular Integral Operators with BMO functions. *ArXiv e-prints*, January 2016.
- [40] Malempati M. Rao and Zhong D. Ren. *Theory of Orlicz spaces*, volume 146 of *Monographs and Textbooks in Pure and Applied Mathematics*. Marcel Dekker, Inc., New York, 1991.
- [41] J. Michael Wilson. Weighted inequalities for the dyadic square function without dyadic  $A_\infty$ . *Duke Math. J.*, 55(1):19–50, 1987.

G. H. IBÁÑEZ-FIRNKORN. FAMAF, UNIVERSIDAD NACIONAL DE CÓRDOBA, CIEM-CONICET & BCAM - BASQUE CENTER FOR APPLIED MATHEMATICS  
*E-mail address:* gonzaibafirn@gmail.com

ISRAEL P. RIVERA-RÍOS. UNIVERSIDAD DEL PAÍS VASCO/EUSKAL HERRIKO UNIBERTSITATEA, DEPARTAMENTO DE MATEMÁTICAS/MATEMATIKA SAILA  
*E-mail address:* petnapet@gmail.com